NASA CONTRACTOR REPORT

NASA CR-350

NASA CR-350

N66-12129

(ACCESSION NUMBER) | (THRU)
243 | 1
(PAGES) | (CODE)
 | 30
(NASA CR OR TMX OR AD NUMBER) | (CATEGORY)

FACILITY FORM 602

GPO PRICE $ ______

CFSTI PRICE(S) $ ______

Hard copy (HC) ______

Microfiche (MF) ______

ff 653 July 65

MINIMAX ATTITUDE CONTROL OF AEROBALLISTIC LAUNCH VEHICLES

Prepared under Contract No. NAS 8-11421 *by*
HUGHES AIRCRAFT COMPANY
Culver City, Calif.
for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • DECEMBER 1965

NASA CR-350

MINIMAX ATTITUDE CONTROL OF AEROBALLISTIC LAUNCH VEHICLES

Prepared under Contract No. NAS 8-11421 by
HUGHES AIRCRAFT COMPANY
Culver City, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $6.00

CONTENTS

INTRODUCTION

An engineer confronted with the problem of designing an autopilot system for a rocket vehicle whose ratio of length to diameter is moderate to high is faced with a host of new problems. These problems arise principally from the fact that rocket structures are usually highly flexible because of the requirement for a low ratio of structural weight to total vehicle weight (usually only 10 percent of the vehicle weight is made up of structure that can resist elastic deformation of the vehicle under the various forces to which it is exposed in flight). For such systems, the control engineer can no longer concern himself only with the pitch, yaw, roll, and translation of the vehicle to obtain the requirements that his control system must satisfy, but rather he must broaden the scope of his analysis to include problems associated with the elastic structure, both to discover the control system requirements and to assure the compatibility of the control system with the dynamic characteristics of the elastic structure. Specifically, if he does not consider the effects of the elastic structure on the control system requirements, he may well design a system that will be unstable when actually installed in the vehicle; if he does not consider the compatibility of the control system with the dynamic characteristics of the elastic structure, he may design a stable control system that will cause structural failure because of dynamic elastic deformations arising from this operation. The above problems are not academic, but very real, and occur almost universally in rocket vehicle design. The basic reason for this is that the highly elastic heavy structure will usually have an elastic response mode whose frequency falls well within the control frequency band and <u>cannot</u> be neglected, in contrast to conventional aircraft design where the control band can usually be chosen well below the important elastic response frequency bands.

One mathematical formulation of a problem within the framework outlined above was stated in the original procurement request, and is in essence repeated here, paraphrased to conform to the particular problem considered.

The steady-state missile dynamics are represented by a homogeneous system of differential equations.

$$\dot{x} = Ax$$

where x is an n-vector and A an $n \times n$ matrix, together with an initial state $x(0) = x^o$. The system expresses various forces and torques due to structural and aerodynamic effects. The time interval $0 \leqq t \leqq T$ for the problem is assumed to be sufficiently short that A is considered constant. A scalar disturbance γ is introduced to represent wind effects in the missile where, analytically, $\gamma \triangleq \gamma(t)$ is restricted to a class of functions Γ. In order to maintain stability under a disturbance as described above, control elements are introduced into the system as a scalar ψ in some class of functions Ψ.

The system is now rewritten to include γ and ψ as

$$\dot{x} = Ax + a\psi + b\gamma \quad , \qquad x(0) = x^o$$

where a and b are constant n-vectors. Let L_i be a given set of constants such that $|x_i| < L_i$, $i = 1, 2, \cdots n$ assures that the system remains stable.

The basic control problem may now be posed as an optimum control one; that is, a control must be found that will ensure $|x_i| \leqq L_i^* < L_i$ for all functions γ in the class where L_i^* is the least such bound for each i. Since the control law which gives L_i^* may be different for each i, it may be necessary to specify constraints on L_i^*. For example, if $L_i^* = cL_i$ where $0 < c < 1$, the control law that minimizes c could be determined. Or the control law that minimizes L_j^* only while satisfying $|x_i| < L_i$, $i = 1 \cdots n$ could be found.

The particular problem formulation for the present investigation is to find ψ that realizes

$$\min_{\psi \in \Psi} \max_{\gamma \in \Gamma} \|x\|$$

Because of the difficulties inherent in this particular version of the minimax problem, a related formulation is used

$$\max_{\gamma \epsilon \Gamma} \min_{\psi \epsilon \Psi} \lim_{n \to \infty} \left\{ \int_{o}^{T} \left[x(t) \cdot Qx(t) \right]^{n} dt \right\}^{1/n}$$

where Q is a n×n positive semidefinite matrix.

The rationale behind the modified form is discussed in Section 2, wherein the problem is specialized to determining ψ for

$$\min_{\psi \epsilon \Psi} \int_{o}^{T} \left[x(t) \cdot Qx(t) \right]^{n} dt \tag{1}$$

which is related to finding ψ for

$$\min_{\psi \epsilon \Psi} \int_{o}^{T} \left[\sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu} \right) \psi_{2\nu}(x) + \frac{1}{2} \psi^{2} \right] dt \tag{2}$$

where $\psi_{2\nu}$ is a positive semi-definite multinomial form. As a first approximation to the minimax problem, criterion (1) is used with n = 1, resulting in a linear controller; the derivation of the design procedure for the linear approximation is presented in Section 3. Better approximations are obtained using criterion (2); these are discussed in Section 4.

The design procedure developed in Section 3 has been programmed for the IBM 7094; the programs are discussed in Section 5, as are the results of applying the design procedure to a five-dimensional model of the booster (taken to be two linked rigid bodies). The controllers thus designed were simulated on the IBM 7094; the results of the simulations are presented in Section 6.

A more complete mathematical model, including the effects of body bending, sloshing, and sensor dynamics, is developed in Section 1. The 26-pole model derived is about as large as the capacity of the computer will allow without modification of the programs, and is also about as small as it could be to be a good representation of the physical system.

A recapitulation of the results of this investigation, along with suggestions for further study, is given in Section 7. In addition, several appendices are included for background material, for detailed derivations, or because they were published as papers based on the material generated under this contract.

1. MATHEMATICAL MODEL

In the analysis or synthesis of any physical system, one of the first and most important steps is the selection of an appropriate mathematical model. For initial synthesis and feasibility analysis, a rather gross approximation to the actual dynamics may suffice; for final analysis and simulation, a more faithful description is usually necessary. When the physical object is as complex as the non-rigid aeroballistic vehicle treated herein, the problem of choosing the model is especially difficult, not necessarily from the point of view of the dynamic description, but rather from that of determining how much fidelity is required to achieve a sufficiently accurate assessment of the behavior.

The model derived in this section is intended to be complete enough for an accurate determination of the dynamics of the vehicle, and yet of a low enough order to allow computer simulation. Included in this model are:

1) Aerodynamic forces (considered to be located at the vehicle center of pressure).
2) Inertia reaction torques on vehicle motion due to nozzle dynamics ("tail wags dog" effect).
3) A flexible vehicle (bending modes).
4) A liquid fuel (sloshing modes).
5) Crosscouplings of bending and sloshing modes with rigid body modes because of engine thrust.
6) Couplings of bending and sloshing modes with engine dynamics.
7) Sensor dynamics.

Not included are the effects of:

1) A distributed aerodynamic force on vehicle motion.
2) Flutter due to aerodynamic forces. (This is an aeroelastic phenomenon, to be accounted for during the airframe structure design phase; it is not a control problem.)
3) Bending motion on aerodynamic forces.

The coordinate system definition and the definition of important physical constants are shown in Figure 1-1. The moving coordinate system is located with its origin at the cg of the booster, oriented as shown. If the vehicle were rigid and the nozzle undeflected, the x-axis and the center line would coincide; the z-axis is in the plane of the local vertical and the velocity vector, which is also assumed to be the plane of the deflected booster centerline. The angle w denotes the rotation of the center line due to a deflection of the nozzle in the absence of bending; in this case the center line is taken as the line through the centers of gravity of the vehicle minus nozzle and of the nozzle alone.

EQUATIONS

The first four equations, which describe the motion of the vehicle, may be obtained through the use of Figure 1-1, by summing forces and moments; Table 1-1 is a list of symbols. (Note that α, ϕ, ψ_g, and β have been assumed small.)

Normal Force Equation.

The normal force equation is found by summing forces in the direction normal to the missile's longitudinal axis (z direction) and equating it to the acceleration in that direction.

$$Ma_z = N_\alpha \alpha + Mg \cos \theta \, \phi + (N_\alpha - D) w - T \psi_g + T_c \beta \tag{1-1}$$

Axial Force

The force along the longitudinal axis of the missile (the x direction) is

$$Ma_x = T_c + T - D - Mg \cos \theta \tag{1-2}$$

External Moment Equation

The moment equation is found by summing the moments about the center of gravity of the missile

$$I\ddot{\phi} = N_\alpha \ell_p \alpha + Dv + N_\alpha \ell_p w - (T + T_c) u_g + T \ell_g \psi_g - T_c \ell_g \beta \tag{1-3}$$

The following kinematical equation will be useful:

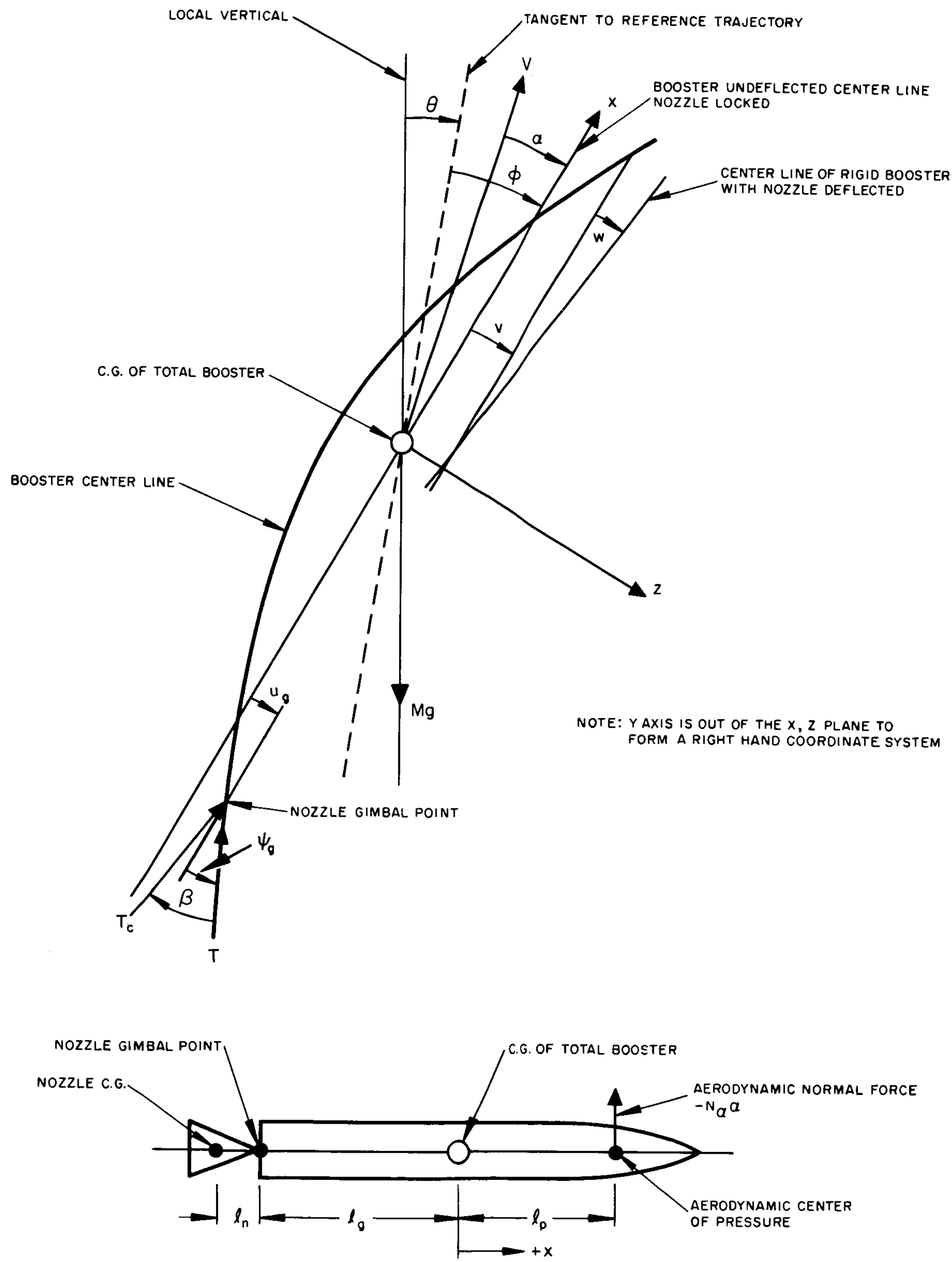


Figure 1-1. Coordinate system for analysis of the flexible booster.

$a_{( )}$	-	acceleration in direction identified by subscript
β_a	-	angle of actuator deflection
ϕ	-	angle between reference trajectory and missile's longitudinal axis
α	-	angle of attack
$\dot{\alpha}_w$	-	crosswind acceleration
β	-	angle of nozzle deflection
q_i	-	normalized deflection of i^{th} bending mode
ζ_i	-	normalized deflection of i^{th} sloshing mode
g	-	acceleration of gravity
N_α	-	normal aerodynamic force per unit of angle of attack
D	-	aerodynamic drag
T	-	thrust of inactive engines
T_c	-	thrust of active engines
V	-	nominal absolute velocity of missile (c. g.)
M	-	total mass of missile
I	-	total moment of inertia of missile
M_n	-	mass of active engines
I_n	-	moment of inertia of active engines
m_i	-	effective mass of fluid for i^{th} sloshing mode
x_i	-	effective location of mass of fluid for i^{th} sloshing mode
ζ_{Si}	-	damping ratio of i^{th} sloshing mode
ω_{Si}	-	natural frequency of i^{th} sloshing mode
ζ_{Bi}	-	damping ratio of i^{th} bending mode
ω_{Bi}	-	natural frequency of i^{th} bending mode

Table 1-1. Nomenclature.

$\lambda_{i(\ )}$	-	normalized slope of i^{th} bending mode at location identified by subscript
$\phi_{i(\ )}$	-	normalized deflection of i^{th} bending mode at location identified by subscript
ℓ_n	-	distance between center of mass of engine and the gimbal point (positive for gimbal point forward of center of mass)
ℓ_g	-	distance between center of mass of missile and the nozzle gimbal point (positive for gimbal point aft of center of mass)
ℓ_p	-	distance between center of mass of missile and center of pressure (positive for c.p. forward of c.g.)
B_n	-	friction damping factor for nozzle travel
K_n	-	effective spring constant between nozzle and case
K_a	-	effective spring constant of actuator arm
θ	-	angle between local vertical and reference trajectory
F_i	-	$\int_{m_i} (r_i)^2\, dm = m_i\,(r_i)^2$ where: r_i is the radius of gyration of the sloshing fluid. The integration is taken over the effective mass of the sloshing fluid (m_i).
v	-	z-axis intercept of the center line of rigid body with nozzle deflected
w	-	rotation of center line of rigid booster due to nozzle deflection
K_s	-	static sensitivity of actuator
τ_s	-	actuator time constant
ω_s	-	actuator natural frequency
ζ_s	-	actuator damping ratio

Table 1-1. Nomenclature (continued).

Normal Acceleration

The vehicle acceleration in the z direction is given by:

$$a_z = -V\dot{\alpha} - a_x \alpha + V\dot{\phi} + V\dot{\alpha}_w \tag{1-4}$$

The remaining equations require some explanation of the representation used for the bending and sloshing dynamics. Both effects result from the motion of continuous media, and the fundamental dynamical equations. Application of the technique of separation of variables to these partial differential equations results in an infinite sequence of sets of total differential equations for each. Each set in the sequence consists of differential equations in spatial coordinates and a differential equation in time; the solution to such a set is called a mode. For a detailed discussion, see the references. The resulting time equations are:

Body Bending Equation

The bending equations are found by applying Hamilton's principle. This yields for, the k^{th} bending mode,

$$\begin{aligned} & M \left[\ddot{q}_k + 2\zeta_{Bk}\,\omega_{Bk}\,\dot{q}_k + (\omega_{Bk})^2 q_k\right] - M\,(a_x + g\cos\theta)\,\lambda_{kg}\,v \\ & - T\phi_{kg}\,\psi_g - \left(\sum_i F_i\,\lambda_{kx_i}\right)(\dot{w} + \ddot{\phi}) + \sum_j \sum_i F_i\,\lambda_{kx_i}\,\lambda_{jx_i}\,\ddot{q}_j \\ & - \sum_\ell m_\ell \left[\phi_{kx_\ell}\,\ddot{\zeta}_\ell - (\lambda_{kx_\ell} + \lambda_{kg})(a_x + g\cos\theta)\,\zeta_\ell\right] \\ & = -(M_n\,\ell_n\,\phi_{kg} + I_n\,\lambda_{kg})\,\ddot{\beta} - T_c\,\phi_{kg}\,\beta \end{aligned} \tag{1-5}$$

Fluid Sloshing Equations

The sloshing equations are obtained by the same method as the bending equations. This method yields for, the k^{th} sloshing mode,

$$\ddot{\zeta}_k + 2\zeta_{Sk}\omega_{Sk}\dot{\zeta}_k + (\omega_{Sk})^2\zeta_k + (a_x + g\cos\theta)\,w + x_k\ddot{w} + \ddot{v} - a_z$$

$$- V\dot{\phi} - x_k\ddot{\phi} + \sum_i (a_x + g\cos\theta)\,\lambda_{ix_k} q_i - \phi_{ix_k}\ddot{q}_i = 0 \qquad (1\text{-}6)$$

The spatial equations may be solved to determine the positions and slopes necessary to solve (1-5) and (1-6); or, more likely, they may be found experimentally.

Figure 1-1 plus the quantities entering into (1-5) and (1-6) lead to the following geometrical equations:

Center Line Deflection

The deflection of the vehicle center line (displacement of c. g.) due to nozzle deflection and sloshing fluid is given by

$$Mv = M_n\ell_n\beta - \sum_j m_j\zeta_j \qquad (1\text{-}7)$$

Center Line Rotation

The rotation of the vehicle center line due to nozzle deflection and sloshing fluid is

$$Iw = -(I_n + M_n\ell_n\ell_g)\,\beta - \sum_j m_j x_j \zeta_j + \sum_i \sum_j \lambda_{jx_i} F_i q_j \qquad (1\text{-}8)$$

Engine Gimbal Point Deflection

The deflection of the engine gimbal point from the undeflected center line of vehicle is given by

$$u_g = v - \ell_g w + \sum_i \phi_{ig} q_i \qquad (1\text{-}9)$$

Engine Gimbal Point Slope

The slope at the engine gimbal point is

$$\psi_g = -w - \sum_i \lambda_{ig} q_i \qquad (1\text{-}10)$$

There is one final dynamic equation:

Engine Dynamics

The engine dynamics are found by summing the moments about the engine gimbal point. This yields

$$I_n \ddot{\beta} + B_n \dot{\beta} + (K_n + K_a)\beta - M_n \ell_n a_z + (I_n + M_n \ell_m \ell_g)\ddot{\phi}$$

$$+ M_n \ell_n g \cos\theta\phi + M(a_x + g\cos\theta)v - I_n \ddot{\psi}_g - M_n \ell_n (a_x + g\cos\theta)\psi_g$$

$$- M_n \ell_n \ddot{u}_g - \sum_j (a_x + g\cos\theta) m_j \zeta_j = K_a \beta_a \tag{1-11}$$

After elimination of the intermediate variables of v, w, u_g, and ψ_g by direct substitution, Equations 1-1 through 1-11 reduce to Equations 1-12 through 1-16 (the coefficients γ are listed in Table 1-2).

Normal Force Equation

$$\boxed{\dot{\alpha} + \gamma_1 \alpha - \dot{\phi} + \gamma_2 \phi + \sum_i (\gamma_3^i q_i) + \sum_j (\gamma_4^j \zeta_j) = \gamma_5 \beta + \dot{\alpha}_w} \tag{1-12}$$

Moment Equation:

$$\boxed{\ddot{\phi} + \gamma_6 \alpha + \sum_i (\gamma_7^i q_i) + \sum_j (\gamma_8^j \zeta_j) = \gamma_9 \beta} \tag{1-13}$$

Nozzle Dynamics Equation:

$$\boxed{\begin{aligned} &\gamma_{10} \ddot{\phi} + \gamma_{11} \alpha + \sum_i (\gamma_{12}^i \ddot{q}_i) + \sum_i (\gamma_{13}^i q_i) + \sum_j (\gamma_{14}^j \ddot{\zeta}_j) \\ &+ \sum_j (\gamma_{15}^j \zeta_j) + \gamma_{16} \ddot{\beta} + \gamma_{17} \dot{\beta} + \gamma_{18} \beta = \gamma_{19} \beta_a \end{aligned}} \tag{1-14}$$

Bending Deflection Equation (k^{th} mode):

$$\gamma_{20}\ddot{\phi} + \sum_i(\gamma_{21}^i \ddot{q}_i) + \ddot{q}_k + \gamma_{22}\dot{q}_k + \sum_i(\gamma_{23}^i q_i) + \gamma_{24} q_k + \sum_j(\gamma_{25}^j \ddot{\zeta}_j) + \sum_j(\gamma_{26}^j \zeta_j) = \gamma_{27}\ddot{\beta} + \gamma_{28}\beta \tag{1-15}$$

Sloshing (k^{th} mode):

$$\gamma_{29}\ddot{\phi} + \gamma_{48}\alpha + \dot{\alpha} + \sum_i(\gamma_{30}^i \ddot{q}_i) + \sum_i(\gamma_{31}^i q_i) + \sum_j(\gamma_{32}^j \ddot{\zeta}_j) + \sum_j(\gamma_{33}^j \zeta_j) + \gamma_{34}\ddot{\zeta}_k + \gamma_{35}\dot{\zeta}_k + \gamma_{36}\zeta_k = \gamma_{37}\beta + \gamma_{38}\ddot{\beta} \tag{1-16}$$

Sensor Equations

In addition to the equations of motion, it is necessary to introduce equations that describe the output of any sensors used. Sensors mounted on the vehicle sense the rigid body motion, the motion of the vehicle arising from the bending modes, the sloshing modes and the engine dynamics. The following three equations show the total inputs to the various sensors. It should be noted that these equations apply only to "perfect" sensors. No sensor dynamics have been included. The addition of sensor dynamics requires the cascading of the sensor dynamics with the output of the "perfect" sensor defined by Equations 1-17, 1-18, and 1-19.

Angular Displacement

The angular displacement sensed by an instrument located at station P along the missile's longitudinal axis is given by

$$\phi_{sensor} = \phi + \sum_i(\gamma_{39}^i q_i) + \sum_j(\gamma_{40}^j \xi_j) + \gamma_{41}\beta \tag{1-17}$$

Rate

The input to an angular rate sensing instrument located at station P along the missile's longitudinal axis is given by

$$\boxed{\phi_{sensor} = \phi + \sum_i \gamma_{39}^i q_i + \sum_j \gamma_{40}^j \xi_j + \gamma_{41}\dot{\beta}} \tag{1-18}$$

Normal Acceleration

The input to an accelerometer which senses acceleration normal to the longitudinal axis of missile located at station A is given by

$$\boxed{\begin{aligned} n_A = {} & \gamma_{42}\ddot{\phi} + \gamma_{43}\dot{\phi} - \gamma_{43}\dot{\alpha} + \gamma_{44}\alpha + \sum_i (\gamma_{45}^i \ddot{q}_i) \\ & + \sum_j (\gamma_{46}^j \ddot{\xi}_j) + \gamma_{47}\ddot{\beta} \end{aligned}} \tag{1-19}$$

The equations governing the sensor dynamics depend upon the mechanization of the sensor; the following are typical sensor dynamics:

Angular Displacement

If it is assumed that the angular orientation is measured by a position gyro, the sensor dynamics can be taken as a pure gain. The attitude reference is a "free" gyro (really a three-axes gimballed gyro), and the orientation of the body relative to the gyro is read out by means of synchro pickoffs. Any dynamics associated with the motion of the gyro would arise as a result of manufacturing inperfections, e.g., bearing torques and mass unbalance. The synchro is essentially a variable transformer; the dynamics associated with the synchro signal arise in the signal processing circuitry and are quite high in frequency. Synchro pickoffs, or analogous linear devices can be used to measure engine gimbal angles.

$$\gamma_1 = \frac{T_c + T - D + N_\alpha}{MV} - \frac{g}{V}\cos\theta$$

$$\gamma_2 = \frac{g}{V}\cos\theta$$

$$\gamma_3^i = -\frac{T}{MV}\lambda_{ig} + \sum_k \left(\frac{F_k}{IV}\lambda_{ix_k}\right)\left(\frac{T - D + N_\alpha}{M}\right)$$

$$\gamma_4^j = -\left(\frac{T - D + N_\alpha}{IMV}\right) m_j x_j$$

$$\gamma_5 = -\frac{T_c}{MV} + \left(\frac{T - D + N_\alpha}{MVI}\right)(I_n + \ell_n \ell_g M_n)$$

$$\gamma_6 = -\frac{N_\alpha \ell_p}{I}$$

$$\gamma_7^i = +\frac{T}{I}(\phi_{ig} + \ell_g \lambda_{ig}) - \frac{(N_\alpha \ell_p + T\ell_g)}{(I)^2}\left(\sum_k F_k \lambda_{ix_k}\right) + \frac{T_c}{I}\phi_{ig}$$

$$\gamma_8^j = -\frac{T_c + T - D}{MI} m_j + \frac{(N_\alpha \ell_p + T\ell_g)}{(I)^2} m_j x_j$$

$$\gamma_9 = -\frac{T_c}{I}\ell_g - \frac{(T + T_c - D)}{MI} M_n \ell_n - \frac{(N_\alpha \ell_p + T_c \ell_g)}{(I)^2}(I_n + \ell_n \ell_g M_n)$$

Table 1-2. Coefficient definitions.

$$\gamma_{10} = \frac{(I_n + \ell_n \ell_g M_n)}{\ell_n M_n}$$

$$\gamma_{11} = \frac{-N_\alpha}{M}$$

$$\gamma_{12}^i = \frac{I_n + \ell_n \ell_g M_n}{IM_n \ell_n} \left(\sum_k F_k \lambda_{ix_k}\right) - \phi_{ig} + \frac{I_n \lambda_{ig}}{M_n \ell_n}$$

$$\gamma_{13}^i = \left(\frac{T_c - N_\alpha}{MI}\right) \left(\sum_k F_k \lambda_{ix_k}\right) + \frac{(T_c - D)}{M} \lambda_{ig}$$

$$\gamma_{14}^j = + \frac{m_j}{M} + \left(\frac{I_n + M_n \ell_n \ell_g}{IM_n \ell_n}\right) m_j x_j$$

$$\gamma_{15}^j = - \frac{(T_c - N_\alpha)}{MI} m_j x_j$$

$$\gamma_{16} = \left[\frac{I_n I - (I_n + \ell_n \ell_g M_n)^2}{IM_n \ell_n}\right] - \frac{M_n \ell_n}{M}$$

$$\gamma_{17} = \frac{B_n}{M_n \ell_n}$$

$$\gamma_{18} = \left(\frac{T - D}{M}\right) + \frac{K_n + K_a}{M_n \ell_n} - \frac{(T_c - N_\alpha)}{M} (I_n + \ell_n \ell_g M_n)$$

$$\gamma_{19} = \frac{K_a}{M_n \ell_n}$$

$$\gamma_{20} = \frac{1}{M}\left[\sum_\ell (F_\ell \lambda_{kx_\ell})\right]$$

Table 1-2. Coefficient definitions. (continued)

$$\gamma_{21}^{i} = -\frac{1}{MI}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{kx_{\ell}}\right)\right]\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{ix_{\ell}}\right)\right]$$

$$+\frac{1}{M}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{ix_{\ell}}\,\lambda_{kx_{\ell}}\right)\right]$$

$$\gamma_{22} = 2\,\zeta_{Bk}\,\omega_{Bk}$$

$$\gamma_{23}^{i} = -\frac{T\phi_{kg}}{M}\,\lambda_{ig} + \frac{T\phi_{kg}}{MI}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{ix_{\ell}}\right)\right]$$

$$\gamma_{24} = (\omega_{Bk})^{2}$$

$$\gamma_{25}^{j} = \frac{m_j\,x_j}{MI}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{kx_{\ell}}\right)\right] - \frac{m_j}{M}\,\phi_{kx_j}$$

$$\gamma_{26}^{j} = \frac{\left(T_c + T - D\right)}{M}\,\frac{m_j}{M}\,\lambda_{kx_j} - \frac{T\phi_{kg}}{MI}\,m_j\,x_j$$

$$\gamma_{27} = \frac{I_n}{M}\,\lambda_{kg} - \frac{M_n\,\ell_n}{M}\,\phi_{kg} - \frac{(I_n + M_n\,\ell_n\,\ell_g)}{MI}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{kx_{\ell}}\right)\right]$$

$$\gamma_{28} = -\frac{T_c}{M}\,\phi_{kg} + \frac{\left(T_c + T - D\right)}{M}\,\frac{M_n\,\ell_n}{M}\,\lambda_{kg} + \left(\frac{I_n + M_n\,\ell_n\,\ell_g}{I}\right)\frac{T\phi_{kg}}{M}$$

$$\gamma_{29} = \frac{x_k}{V}$$

$$\gamma_{30}^{i} = -\frac{\phi_{ix_k}}{V} - x_k\,\frac{1}{IV}\left[\sum_{\ell}\left(F_{\ell}\,\lambda_{ix_{\ell}}\right)\right]$$

Table 1-2. Coefficient definitions. (continued)

$$Y_{42} = \ell_A$$

$$Y_{43} = V$$

$$Y_{44} = -\left[\frac{T_c + T - D}{M} - g \cos\theta\right]$$

$$Y_{45}^{i} = \frac{\ell_A}{I}\left[\sum_{\ell}\left(F_\ell \lambda_{ix_\ell}\right)\right] - \phi_{iA}$$

$$Y_{46}^{j} = \frac{m_j}{M} - \frac{\ell_A}{I} m_j x_j$$

$$Y_{47} = \frac{\ell_n M_n}{M} - \frac{\ell_A (I_n + \ell_n \ell_g M_n)}{I}$$

$$Y_{48} = \frac{a_x}{V}$$

Table 1-2. Coefficient definitions. (continued)

Angular Rate

Vehicle angular rates are usually measured by single-axis spring-restrained gyroscopes (rate gyros) which exhibit damped oscillatory behavior (second order). Accelerations about the axis perpendicular to the spin axis and the sensitive axis appear as rate errors, as do cross-coupling effects at large rates. The engine gimbal rates can be measured with a tachometer with negligible dynamics.

Normal Acceleration

Normal acceleration is also usually sensed by devices exhibiting second order behavior, e.g., damped spring-mass linear accelerometers and integrating-gyro accelerometers.

Actuator Dynamics

If the signal from the controller is used to displace the control valve of a hydraulic actuator, the transfer function between the control signal and the actuator displacement may be taken to be of the form

$$\frac{K_s}{(\tau_s s+1)\left(\frac{s^2}{\omega_s^2} + \frac{2\zeta_s}{\omega_s} s+1\right)} \tag{1-20}$$

Model Dynamics

Before the model is completely specified, the number of bending modes and sloshing modes to be included must be determined. Since the natural frequencies of the dynamics increase as the mode number increases, the number of modes is usually determined by neglecting those modes corresponding to natural frequencies "sufficiently high" in comparison with the control system response characteristics. For the nonrigid aeroballistic booster studied herein, four bending modes and three sloshing modes were retained, resulting in a 26-pole transfer function (or 26 state variables). A descriptive block diagram of this model is shown in Figure 1-2. The summation points should be interpreted as indicating that the given input affects the block in question, but not necessarily in the strictly additive manner shown.

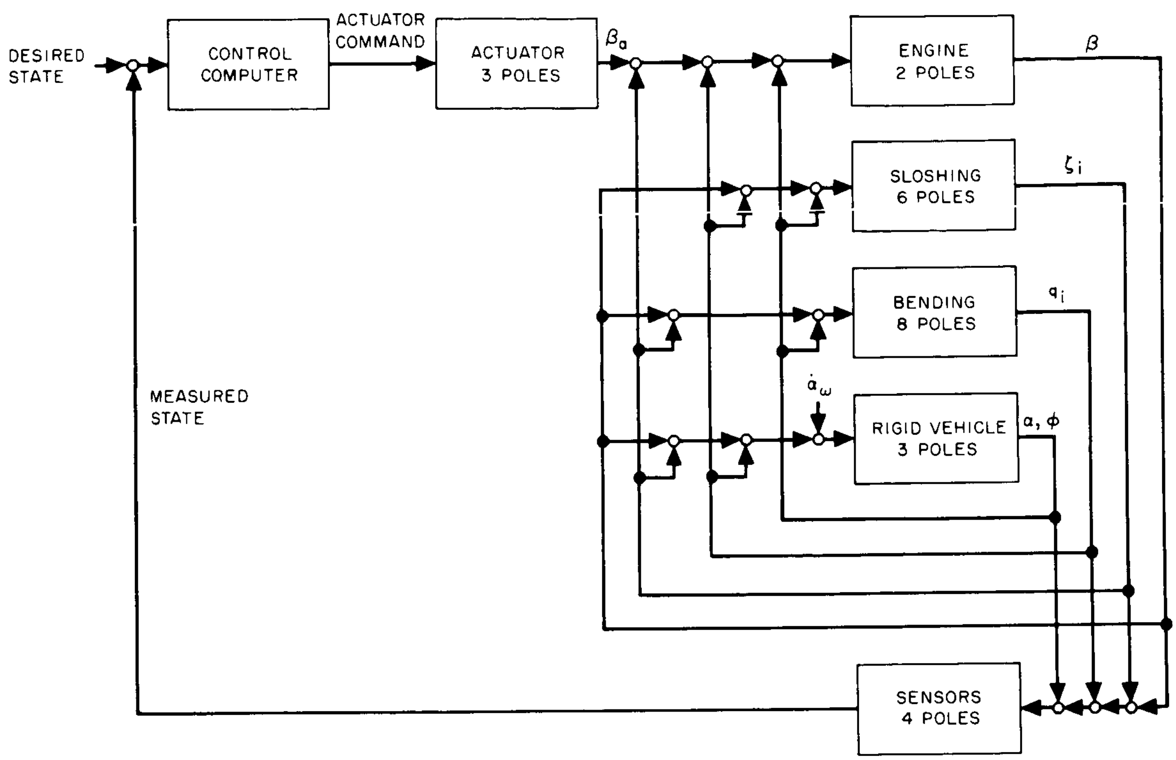


Figure 1-2. Block diagram of 26-pole model.

2. MINIMAX PERFORMANCE CRITERIA

In the original proposal to MSFC, Hughes Aircraft Company defined a class of performance criteria which were to be investigated to determine their usefulness for generating control laws. These criteria are designated "minimax" and pertain to the keeping of specified combinations of states of the given system as small as possible when the system is acted upon by the "worst" of a class of external disturbances. This class of performance criteria can be better described in a mathematical form.

Let the system under consideration be governed by

$$\dot{x} = Ax + a\psi + b\gamma$$

where

A is an n×n plant matrix

a is an n-vector, coupling the control into the system

b is an n-vector, coupling the disturbance into the system

ψ is the scalar control to be chosen

γ is the disturbance.

Let $\gamma = \gamma(t)$ be a member of the set of allowable disturbances Γ, denoted by $\gamma \epsilon \Gamma$; similarly let $\psi \triangleq \psi(x) \epsilon \Psi$, the allowable set of controls. Then one form of minimax criterion is given by the following

$$\min_{\psi \epsilon \Psi} \max_{\gamma \epsilon \Gamma} \|x\| ,$$

where a variety of norms may be used for $\|\cdot\|$. One convenient norm is

$$\|x\| \triangleq \max_{t \epsilon [o,T]} \|x(t)\|_Q$$

where

$$\|x(t)\|_Q \triangleq x(t) \cdot Qx(t)$$

with

$Q = Q^* =$ n×n positive semi-definite matrix.

Using these definitions and the fact that

$$\max_{t \in [o,T]} |f(t)| = \lim_{n \to \infty} \left\{ \int_o^T |f(t)|^n \, dt \right\}^{1/n} ,$$

we can write this minimax criterion as

$$\min_{\psi \in \Psi} \max_{\gamma \in \Gamma} \lim_{n \to \infty} \left\{ \int_o^T [x(t) \cdot Qx(t)]^n \, dt \right\}^{1/n} .$$

Another version of minimax criterion results when one interchanges $\max_{\gamma \in \Gamma}$ and $\min_{\psi \in \Psi}$. Doing this we may write

$$\max_{\gamma \in \Gamma} \min_{\psi \in \Psi} \lim_{n \to \infty} \left\{ \int_o^T [x(t) \cdot Qx(t)]^n \, dt \right\}^{1/n} .$$

This form is mathematically more tractable than the previous one and is the one which Hughes Aircraft Company studied extensively during the past year. It should be emphasized that, in general, one does not get the same value for the performances indices in the two cases. The exact conditions under which such an operation yields the same numerical results before and after the exchange is not known. This is a current area of research both at Hughes Aircraft Company and many other institutions. In what follows we shall only concern ourselves with performance criterion

$$\max_{\gamma \in \Gamma} \min_{\psi \in \Psi} \lim_{n \to \infty} \left\{ \int_o^T [x(t) \cdot Qx(t)]^n \, dt \right\}^{1/n} .$$

Specifically first consider

$$\min_{\psi \in \Psi} \lim_{n \to \infty} \left\{ \int_o^T [x(t) \cdot Qx(t)]^n \, dt \right\}^{1/n} .$$

This in itself is a difficult mathematical problem and must be further simplified before a meaningful solution can be presented. With this in mind we propose studying the minimization of

$$\left\{\int_0^T [x(t)\cdot Qx(t)]^n \, dt\right\}^{1/n}$$

for some large fixed integer n, instead of minimizing

$$\lim_{n\to\infty}\left\{\int_0^T [x(t)\cdot Qx(t)]^n \, dt\right\}^{1/n} .$$

Again rigorous justification of the closeness of these two problems is difficult, but the results of simulations have shown the simplification to be valid. The actual problem considered is

$$\min_{\psi\in\Psi}\left\{\int_0^T [x(t)\cdot Qx(t)]^n \, dt\right\}^{1/n} .$$

The first approximation to be considered is the case when $n = 1$. This is described in Section 3 and the necessary background is given in Appendix C. For the case $n = 1$ an exact solution may be found, but in the case $n > 1$ only an approximate solution is easily obtained.

In Section 4 minimization of performance criteria of the form

$$\int_0^T \left[\sum_{\nu=1}^{\infty}\left(\frac{1}{2\nu}\right)\psi_{2\nu}(x) + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^2\right] dt$$

where $\psi_{2\nu}(x)$ is a positive semi-definite multinomial form, is described. It should be noted that the performance criteria

$$\int_0^T [x(t)\cdot Qx(t)]^n \, dt$$

can be put in this form by letting $\psi_{2n}(x) = (x\cdot Qx)^n$ and $\psi_{2\nu}(x) = 0$ when $\nu \neq n$.

Thus the problems considered there are of a more general nature than minimization of

$$\int_0^T [x(t)\cdot Q\,x(t)]^n \, dt \quad .$$

3. LINEAR DESIGN PROCEDURE FOR MINIMAX PROBLEM

Using the standard techniques of control system analysis and design such as root locus, Bode analysis, etc., for high order systems usually results in a great deal of successive approximation and often depends heavily on the ingenuity of the particular investigator carrying out the analysis. Indeed, the concept of optimum design is not even considered in general, for the very nature of the methods used necessitates individual attention for each problem. These difficulties, coupled with the fact that the "classical" methods have no natural extension to systems with time-varying parameters and nonlinearities, have led modern contributors in control theory to consider the problem from a fresh viewpoint — that of state space analysis. The advantages of this approach are manifold. Involved and often very specialized computations are reduced to common matrix manipulations quite amenable to present day high speed computing devices. The idea of optimal design can be stated in a very simple manner which is applicable to a myriad of problems. Physical variables are not lost in a jungle of mathematical manipulation, but rather maintain their identity throughout the analysis of a problem, thus permitting new insight into the role of these variables in the overall design. Perhaps the most important single advantage of state analysis is that there is no conceptual difference in the presentation of linear, linear time-varying, or nonlinear problems. All this has inspired a vast amount of research which, in turn, has resulted in a flood of publications in the area. In their zeal to contribute, many writers have overlooked completeness in favor of broadness, and even though many complex problems have been considered, efficient engineering solutions to many important control problems are lacking. Although the work of R. E. Kalman is nearly definitive from a mathematical viewpoint, it neglects the design problem from the point of view of the practicing control engineer. Because of his determination to solve the autonomous and nonautonomous optimal control problems simultaneously, Kalman failed to attain ultimate simplicity in the solution of the former problem. Thus his theory for constant coefficient systems depends on the steady-state solution of a matrix differential equation, a numerically cumbersome proposition. Work at Hughes, on the other hand, as presented in the paper, "High Order System Design Via State Space Considerations,"* has permitted the design of optimal single channel controls by purely algebraic means, thus reducing computer time and allowing for extension to high order systems. The following discussion is based primarily on that paper.

In general, a linear system with single channel control can be represented by the set of differential equations

$$\dot{x} = Ax + a\psi \tag{3-1}$$

where x is the state vector, A is the matrix of the plant parameters, a is an actuator vector, and ψ is a scalar control function, assumed here to be a linear combination of the states at any instant of time. The object of the design procedure under consideration is to find this linear feedback relationship so as to optimize the performance of the resulting "closed loop" system. In particular, the elements of a vector g are sought such that $\psi = g \cdot x$ minimizes an integral of the form

$$\Phi = \frac{1}{2}\int_0^\infty (x \cdot Cx + \psi^2)dt \tag{3-2}$$

where C is a symmetric non-negative definite matrix. The choice of the matrix C is equivalent to specifying the nature of the optimality to be considered. Indeed, it directly determines the performance of the resulting system. This matrix can be appropriately chosen only in the context of a particular problem. For aerospace vehicle stabilization, quantities such as structural load, pitch error, etc., must be kept below certain bounds while the maximum of some critical quantity such as lateral drift is minimized. Indeed, load, pitch error, and drift can be expressed as a linear combination of state variables of the form

$$\left| q^i \cdot x \right|, \qquad i = 1, 2, \cdots, \tilde{m} . \tag{3-3}$$

By noting that in the integral

$$\int_0^\infty \left[(q^i \cdot x)/\kappa_o \right]^2 dt \tag{3-4}$$

the total contribution of time at which $\left|q^i \cdot x\right| > \kappa_o$ holds is "penalized" disproportionately compared to the times at which $\left|q^i \cdot x\right| < \kappa_o$ holds,

*Presented at 1965 JACC.

the matrix C can be specified. In particular, the abovementioned objectives are satisfied to a first approximation if

$$C = \kappa_1 q^1 (q^1)^* + \kappa_2 q^2 (q^2)^* + \cdots + \kappa_{\tilde{m}} q^{\tilde{m}} (q^{\tilde{m}}) \qquad (3\text{-}5)$$

In the analysis of systems by state space approaches, the concepts of controllability and observability lie at the very foundation of system theory. Mathematically, if controllability is ensured, there always exists some control to bring the system from one arbitrary point in state space to any other. If an output is observable, that output cannot vanish identically for a finite period of time unless the system is totally at rest. (See Appendix F for mathematical criterion for linear systems.) Heuristically, the lack of controllability implies an "open circuit" somewhere in the system input (i. e., one or more modes of the system cannot be reached). Analogously, if an output is not observable, then an "open circuit" exists in that output path. These ideas do not appear in the classical transfer function methods, for the transfer function itself is only a valid representation of the observable and controllable part of a linear dynamic system. Appropriate tests for these criteria can be made only when the system is represented in the "natural" form of simultaneous differential equations in many variables.

The theoretical development of the actual design method in question is fairly involved and so only an outline of the ideas will be discussed here. Basically, Pontriagin's Maximum Principle is applied to the system (3-1) to give the necessary conditions for minimizing the integral (3-2). Combining the resulting equations with stability requirements as formulated by Liapunov's Second Method results in a unique feedback law which can always be found by purely algebraic operations. In particular, the design method suggested is based on finding the resolvent of the matrix A (evaluating the matrix of polynomials in s given by $\left[sI - A\right]^{-1}$). This can be accomplished directly by Leverrier's algorithm or indirectly by standard matrix manipulations. In finding the resolvent, the characteristic equation for the system is also explicitly displayed. From these quantities, and the chosen performance matrix C, the characteristic equation for the optimal closed loop system can be directly determined. Then, with the aid of a specially derived relationship between the open and closed looped systems the vector g can be found. It is to be noted that thus far system controllability has been assumed. The "amount" of controllability as given by a specific algebraic criterion determines the magnitude of the vector g and hence the feasibility of practical implementation of linear control.

In addition to being optimal in the sense discussed previously, systems designed by these methods are remarkably adaptive to large variations or to saturations in both feedback signals and actuator characteristics. It is shown in Appendix C that once a system is designed to minimize a quadratic performance index, a Liapunov function can be found which will guarantee the stability of the system to certain initial perturbations for considerable variations in the feedback signal. Furthermore, under these conditions the modified performance index

$$1/2\int_0^\infty \left[x \cdot Cx + (g \cdot x)^2 \right] dt \tag{3-6}$$

will not be increased beyond the nominal minimum value found for the system if no perturbations were present. Adaptivity to feedback saturation is particularly important when the elements of the control vector g are relatively large.

In a wide class of problems, saturation may not be permissible or it may be very desirable to keep the magnitude of the control vector small. This can be readily accomplished by taking the matrix C in (3-2) to be identically zero. Minimizing the resulting performance criterion will then be equivalent to minimizing the "amount" of feedback in a least-squares sense. With a stable plant this so called "minimum effort" control reduces trivially to no control at all. However, with an unstable plant such a criterion generates a closed loop system whose poles consist of the stable plant poles and the reflections in the imaginary axis of the unstable ones. In this case then, the optimal control law can be tested easily and compared with other criteria.

In the design of large aeroballistic launch vehicles there are situations in which linear feedback can yield an exact answer to the minimax control problem. This occurs when an arbitrary linear combination of state variables can be forced to decay directly from an initial perturbation. The necessary and sufficient conditions for the existence of such an "ultraminimax" control are shown in Appendix G along with an explicit formula for that control when it exists. It is also shown there that ultraminimax control is a limiting case minimizing a criterion of the form (3-2) when the terms in the state variables are increasingly weighted in comparison to the control term.

It is to be emphasized that designs discussed up to this point are really incomplete, for they assume measurement and feedback of all state variables — a highly unlikely situation in common problems. To supplement these methods, a filter has been designed which can approximately generate $g \cdot x$ given only incomplete state information. That is, the output of linear sensors measuring independent observable variables can be used to generate the optimal feedback control. The filter configuration in general consists of parallel networks, each operating on a specific sensor output. These networks have common poles which are completely arbitrary. The number of poles needed in the filter is usually equal to, or slightly greater than, the quantity $n/m - 1$, but less than the quantity $n - m$ when n is the order of the given system and m is the number of independent sensor outputs. It is found that the optimal system poles as determined above are mechanized in the closed loop system when the filter is introduced. Furthermore, as the real parts of the chosen filter poles become more and more negative, the corresponding poles of the overall closed loop system approach more and more closely the chosen poles themselves. It is noted that this method of filter design may be limited because of increased sensitivity to sensor noise and system parameter variations when fast poles are introduced. The important problem of supplementing this design with a method for optimizing filter pole locations in accordance with a minimal variance scheme has not yet been considered and is an open research problem. One possibility would be to estimate all state variables via a Kalman filter and then reconstruct $g \cdot x$. This, however, always introduces n new poles to the system. For high order systems, a Kalman filter can be quite cumbersome and expensive. As an approximation, perhaps the dominant poles of the Kalman filter can be used in optimizing the poles of the parallel filter described above. Additional research is needed to determine the feasibility of this proposition.

4. DESIGN PROCEDURE FOR NONLINEAR APPROXIMATION OF MINIMAX PROBLEM

The problem of minimizing a quadratic performance criterion for a linear time-invariant plant subject to a mean square constraint on the cost of control leads to a linear control law. As a natural generalization of this problem, the problem of minimizing a quartic or higher order performance criterion subject to a mean-square constraint on the control may be considered. The theoretical details of the solution to this problem are presented in Appendix E. In it, the known fact that it is necessary to introduce quadratic terms in the performance criterion to ensure stability of the overall system is reviewed. Then the quartic and higher-order terms are introduced in order to impose bounds on specified state variables. In other words, the linear control derived from the quadratic terms stabilizes, and the cubic or higher-order control derived from the quartic or higher-order terms limits, the transient response to the desired region. This is exactly what might be suspected, since the stability of a nonlinear autonomous system in a neighborhood of the origin is determined by the linear terms. The nonlinear control law is derived by finding the unique solution to the Hamilton-Jacobi partial-differential equation for the problem posed.

As shown in Section 2, performance indices of the form

$$\int_0^T \sum_{\nu=1}^{\infty} \left[\left(\frac{1}{2\nu}\right)\psi_{2\nu}(x) + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^2\right] dt \tag{4-1}$$

where $\psi_{2\nu}(x)$ is a $2\nu^{th}$ order positive semidefinite form, ψ is the control to be chosen, and $\psi_{n\ell}$ is the nonlinear portion of the control, arise quite naturally as approximations to a minimax performance index.

This performance index may be interpreted in an alternate manner; namely, minimize

$$\int_0^T \left[\sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \psi_{2\nu}(x)\right] dt \tag{4-2}$$

subject to

$$\int_0^T \psi^2 \, dt \le \rho_1 \tag{4-2b}$$

and

$$\int_0^T \psi_{n\ell}^2 \, dt \le \rho_2 \tag{4-2b}$$

By the use of Lagrange multipliers, one can change (4-2) to (4-1). This is explained in detail in Appendix E.

During the past year, Hughes has studied (4-1) in detail to dtermine an optimal control law which will lead to a minimax type of response for the closed loop system. A concise statement of the problem, as well as the synthesis procedure used, follows below. Minimize the performance index Φ

$$\Phi = \int_0^T \left[\sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \psi_{2\nu}(x) + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^2 \right] dt \tag{4-3a}$$

where

$$\psi_{2\nu}(x) \ge 0, \quad x \ne 0 \tag{4-3b}$$

and

$$\psi_{2\nu}(\mu x) = \mu^{2\nu} \psi_{2\nu}(x) \tag{4-3c}$$

subject to

$$\dot{x} = Ax + a\psi \tag{4-3d}$$

The $\psi_{2\nu}(x)$ are the given $2\nu^{th}$ order forms and ψ is to be found.

The control law

$$\psi(x) = -a \cdot \text{grad}_{(x)} V(x) \tag{4-4a}$$

where

$$V(x) \triangleq \sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \varphi_{2\nu}(x), \tag{4-4b}$$

is the optimal control law for (4-3).

The expressions $\phi_{2\nu}(x)$ must be related to the given expressions $\psi_{2\nu}(x)$ in order to complete the solution. This relationship is separated into two parts characterized by $\nu = 1$ and $\nu > 1$. Considering the case $\nu = 1$ and introducing the notation

$$\psi_2(x) \triangleq x \cdot Cx \tag{4-5a}$$

and

$$\phi_2(x) \triangleq x \cdot Bx \tag{4-5b}$$

it is necessary that B satisfy the equation

$$A^* B + BA - Baa^* B = -C \tag{4-6}$$

For the case $\nu > 1$, it is necessary that the $\psi_{2\nu}(x)$ satisfy the equations

$$\tilde{A}x \cdot \operatorname{grad}_{(x)} \phi_{2\nu}(x) = -\psi_{2\nu}(x) \tag{4-7a}$$

where

$$\tilde{A} \triangleq A - aa^* B \tag{4-7b}$$

Here $\tilde{A}$ corresponds to the stabilized linear portion of the system. With the above relationships, the following holds for the optimal closed loop system.

$$\int_0^T \left[\sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \psi_{2\nu} + \frac{1}{2}\psi^2 + \frac{1}{2}\psi^2_{n\ell} \right] dt = \sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \phi_{2\nu}(x^o) \tag{4-8}$$

Note that there is equality in equation 4-8 and the right hand side is identical to the Liapunov function $V(x^o)$ chosen for the closed loop system. Thus $V(x^o)$ is the unique solution to the Hamilton-Jacobi equation for the problem stated. Optimal control laws found by this procedure may actually saturate when mechanized. The resulting system may become unstable for a large initial condition because of the "limited amount" of control available. Estimates for the allowable range of initial conditions in such a situation are derived for the linear case in Appendix C and for the nonlinear case in Appendix E.

These estimates determine the domain of asymptotic stability of the resulting system. However, when disturbances are coupled into the system, one is faced with a problem of determining the region which bounds the system motion when acted upon by persistent disturbances. Thus it is necessary to find a suitable Liapunov type function for the nonautonomous disturbed system. (This is called "practical stability" by Lasalle and Lefschetz.) Some results have been obtained in this area so far and there is considerable evidence that further results are possible.

It should be mentioned that this problem of determining the "operating" region of a disturbed system results from the fact that the optimal control law was found for an undisturbed system with an initial condition. However, results of Potter and Tung seem to indicate that a system designed on this basis will be the best system when the actual disturbance is white noise.

To gain some insight into the problem of determining the domain of stability of the disturbed system consider Figure 4-1.

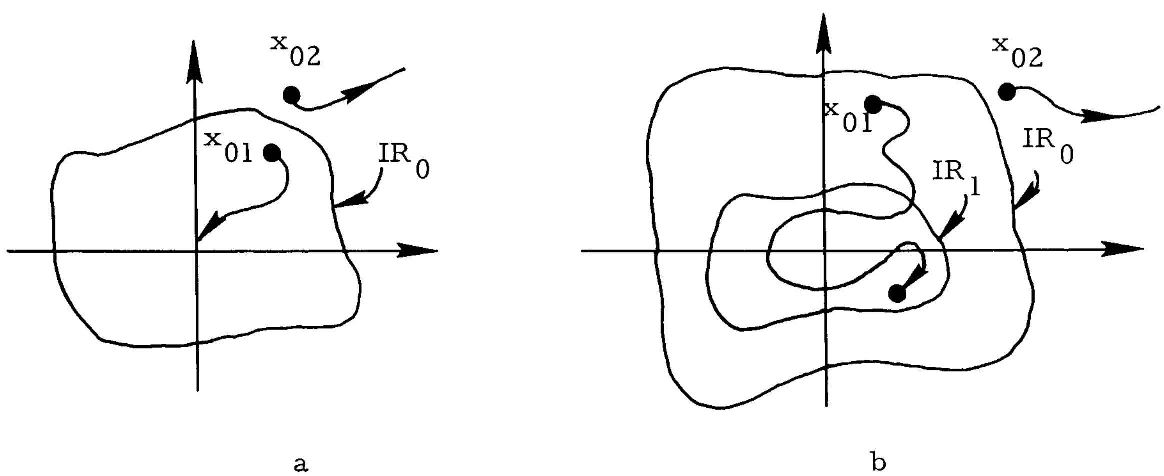


Figure 4-1.

Figure 4-1a depicts the motions of the undisturbed system for two initial conditions x_{01}, and x_{02}. When the system motion starts in the region IR_0, as is the case for the initial condition x_{01}, it eventually returns to the origin. Thus IR_0 represents the domain of asymptotic stability. It should be remarked that IR_0 is bounded when the control is bounded. Theorem 2 of Appendix E characterizes the region IR_0.

Figure 4-1b represents the type of motion that would be expected for a disturbed system when the magnitude of the disturbance is bounded. Any initial condition $x_{01} \epsilon IR_0 - IR_1$ leads to a motion which eventually enters IR_1 and remains there. Any initial condition x_{02} outside of R_0 leads to motion which is unstable because of the assumed boundedness of the available control. Determining the regions IR_1 and IR_0 of Figure 4-1b would allow the designer to assess accurate bounds on motion of the controlled system. This is an area of study which can yield valuable results for designing minimax control systems.

5. COMPUTER PROGRAM FOR LINEAR DESIGN PROCEDURE

BACKGROUND

In this section, the paper entitled "High Order System Design via State-Space Considerations" is mechanized by two basic digital computer programs, including several special subroutines. Included are the listings of the routines, associated diagrams, and a complete dictionary of symbols. The actual IBM 7094 printout of a case using a fifth order Saturn model is given.

The digital programs are referred to by their Fortran language names, CNTRL2 and FILTER. Following is a brief description of their functions: Given the system matrix A, the actuator vector, a, and desired performance index matrix, C, CNTRL2 computes the optimal closed loop poles, then computes the corresponding control vector, g. FILTER, using data from CNTRL2, computes the parameters of a simple multiport filter to approximate the desired result. Another program, CNTRL1, was written which would compute the control vector to achieve arbitrarily specified closed loop poles. This program has been dropped, since the same result could easily be gained by reading in the arbitrary poles as data and bypassing the optimal pole computing portion of CNTRL2.

Several subroutines were written to be called by one or both of the main programs. ALPHAS uses double precision arithmetic in the Leverrier algorithm to compute the coefficients of the characteristic equation of the system matrix. It is used twice in CNTRL2, first to compute the open loop coefficients, then to compute the closed loop coefficients as a check on the main program. POLYWG determines the optimal closed loop poles to be obtained by CNTRL2. SYNTH1 performs most of the vector and matrix manipulations of CNTRL2 to arrive at the control vector, g. OBSMAT computes the observability matrix and filter arrangement matrix for FILTER. Three minor subroutines were written and used by several other routines. MATMPY multiplies matrices, DMATMP does the same in double precision. ORDINV reverses the order of a one-dimensional array, useful since the library subroutines use data which is in reverse order to that required for clarity in these programs.

The programs are completely self-checking, with the exception of POLYWG. To check the validity of POLYWG, an arbitrary example was computed by the use of the Leverrier algorithm, which is itself self-checking. Enough confidence in POLYWG has been developed so that the possible incorporation of the Leverrier algorithm as a check on POLYWG (mentioned in an earlier report) has not been carried out.

Four library subroutines available at Hughes were incorporated into the routines. Their functions are described briefly in this section. No details seem necessary, because anyone outside of Hughes Aircraft wishing to make use of the Linear Design Procedure would have to substitute other subroutines which are at his disposal.

An effort was made to make the programs as self-explanatory as possible, with the aid of many comments in the listings, accompanying block diagrams, a complete dictionary, and annotated comments on the printed output. Hence the explanations in the next section are not all complete in themselves, but serve merely to clarify a few details.

PROGRAM DETAILS

CNTRL2

By having two performance matrices available, it is possible to vary the emphasis on different requirements by varying two weighting factors. In this case, the first performance matrix, C1, was computed to minimize drift, and the second, C2, was computed to minimize load.

A value of either 1 or 2 for MPRNTS must be read in as data, according to whether or not it is desired to write the S-matrices during the program. They should be written if the system transfer functions are desired. If MPRNTS = 2, ENORM, the sum of the absolute values of the S_o elements, is written instead. This is a sufficient indication of roundoff error and takes much less space.

Although the calling statement for ALPHAS contains ELINV in the argument both times, there is no meaning in the second call. ELINV is used only because it is no longer used by the program, hence is a convenient matrix to fill a space.

The coefficients of the optimal closed loop characteristic equation (ALFOPT), as determined by POLYWG, are in inverse order — the standard order of the coefficients being assumed the order of powers of s to which they correspond. The order is reversed and the new array is called OPTALF.

FILTER

The input of the sensor matrix is arranged so that various numbers of sensors can be tried without rewriting the program. Each card of a set of five represents a sensor. A blank or zero card indicates no sensor. The order of the sensors is not important, since the program eliminates the zero rows and labels the remaining ones as Sensor No. 1, Sensor No. 2, etc. The filter numerators are correspondingly numbered.

The program specifies that N values (complex) be read in for POLES (I). These are the arbitrarily chosen filter poles. They should be in order of their desirability, since the program, after determining how many filter poles are required, takes as many as it needs, beginning from the top of the list. Actually no more than n - 1 could be used under any circumstance, so that the n^{th} space could be left blank.

The matrix VK is substituted for OBSERV in order to save OBSERV, since the library subroutine MATS destroys input information.

The GAMMA matrix is obtained by separating the parts of the n-vector d and placing them in adjoining columns as shown in the definitions following Equation (31).* This is done with the aid of the IJDLTL (filter arrangement) matrix which is described further under OBSMAT.

There may seem to be some confusion concerning the H and HSTAR matrices. Note that they share storage locations by means of the EQUIVALENCE statement. Originally read in as HSTAR in order to represent each sensor by a single data card, the rows and columns are then interchanged to be used in computations. When HSTAR is written after the transposition, it is done by writing H but reversing the indices in the WRITE statement.

*All equation numbers in this section are those of the equations in Appendix C, "High Order Design via State-Space Considerations."

Subroutine ALPHAS

ALPHAS mechanizes the recursive relationship expressed by Equations (7a) and (7b).

The single-precision inputs are replaced by double-precision variables. The computations are performed in double precision and the results are replaced by single precision variables for writing or storage. The double-precision dummies are dimensioned internal to the subroutine.

INDEX(NN) is carried along only as a convenience in printout of the ALPHA's and S-matrices. It allows the index to vary from 0 to N, not allowed directly. This feature is probably more confusing than necessary.

Subroutine POLYWG

The coefficient of each even power of s from s^2 through s^{2n} of the polynomial $\Delta_{2n}(s)$ is computed in turn. Coefficients of the odd terms are zero. Beginning with the working dummy COEFF(I) set equal to zero, the subroutine adds on the various parts as expressed in Equation (57). The index is then increased and COEFF(I+1) is computed, etc. Special cases (s^0 and s^{2n}) are computed separately afterward.

Signs of the various terms are determined by the variables SIGNI or SIGNJ which are, at the proper times, either +1 or -1 .

The intermediate result is an n^{th} order polynomial in s^2. The n roots of that polynomial are determined by a root-finding routine. The complex square root of each root in s^2 is taken, giving the 2n roots, half of which are in the left half plane. The complex square root function, CSQRT, provides only one of each complex pair. POLYWG accepts those which are in the left half plane, and changes the sign on those which are not. The resulting roots are the desired optimal roots.

An effort has been made to make the Fortran symbols correspond very closely with the Equations (44) and (57).

Subroutine SYNTH1

Equations (14), (15b) and (20) are mechanized by this subroutine. SYNTH1 first computes the transpose of the controllability matrix, DSTAR.

If it is singular, the control vector cannot be computed so the subroutine writes "System not controllable · · · " . Otherwise GLTL is computed and the subroutine makes a normal return to the main program, CNTRL2.

Note that the equivalence statement is merely a comment, indicating that the actual equivalence statement must appear in the main program. AAT and BLTL, EN and AB are equivalenced as required by library subroutines.

In order to preserve DSTAR for writing in the main program, another matrix, AAT, is substituted to make use of the simultaneous equation subroutine SIMEQ.

Subroutine OBSMAT

This subroutine follows the procedure outlined on page 4 of the reference to compute the observability matrix K (OBSERV). IJDLTL, the filter arrangement matrix, is constructed during the testing of columns for independence. Elements are made 1 if the column is independent, or zero if it is not. After FILTER computes d, the successive components will be placed columnwise in the elements of GAMMA if the corresponding elements IJDLTL are 1. Elements of GAMMA are made zero if the corresponding element of IJDLTL is zero. The dimensions of the GAMMA and IJDLTL matrices are identical.

For computational purposes, a slight refinement has been made on the Gram-Schmidt orthogonalization procedure. Normally, each new column to be adjoined is tested for independence by formation of the orthogonal vector W(I). The column is independent if W(I) is not zero. In OBSMAT a tolerance has been introduced. The magnitude of W (WMAG) is required to be less than 10^{-6} times the average magnitude of previous columns (VMAVM6).

Library Subroutines

A brief functional description of library subroutines used in the Automatic Design Procedure which are available on tape at Hughes Aircraft Company follows.

MATS. Solves simultaneous equations in the form

$$(A)(X) = (B) \quad .$$

(A) and (B) must first be adjoined into an n by n+1 matrix $(A') = (A|B)$. The solution vector (X) is dimensioned separately. The input matrix (A') is destroyed during computation.

SIMEQ. Similar to MATS except (A) is not adjoined to (B) prior to use. Both (A) and (B) are destroyed during computation, with the solution vector (X) remaining as the first column of (A). For this reason (X) must appear equivalenced to (A).

ROOT1. A root-finding routine usually good to $N = 20$. $N + 1$ coefficients of a polynomial in s are input, in order of descending powers of s. As an aid to the subroutine, an initial guess of the roots (APPROX) is made. If the initial guess is zero, the subroutine makes its own first guess.

POLCO. Given the roots of a polynomial, this subroutine gives the coefficients of powers of s, in descending order.

NUMERICAL RESULTS

Although the example shown was chosen arbitrarily to demonstrate the design procedure, it merits some explanation, especially with regard to the filter.

While the design procedure guarantees physical realizability, it does not ensure practicality, as shown by the example. There are two reasons for this: (1) the large performance index weighting factor requires very large feedback gains; (2) since $\dot{\phi}$ is not measured in this example, it must be derived synthetically. In contrast, the filter designed in the Monthly Progress Report dated 15 March 1965 was much more realistic, where $\dot{\phi}$ was measured and the feedback gains were much less.

MAIN PROGRAM
CNTRL 2

READ { MPRNTS, A(I, J), ALTL (I, 1), C1 (I, J), C2 (I, J)

WRITE program title and input data

Find coefficients of open loop characteristic equation via SUBR . ALPHAS. Compute roots. WRITE ALPHA(I) and OLROOT(I). Also write ELINV for later use in FILTER program.

READ Performance Weighting Factors CAPPA1 & CAPPA2

$C = K_1 C1 + K_2 C_2$

WRITE CAPPA1, CAPPA2 & C(I, J)

Compute optimal coefficients (OPTALF) of closed loop characteristic equation by POLYWG.

Compute control vector GLTL by SYNTH1. DSTAR, the transpose of the controllability matrix, is computed in the process. WRITE DSTAR

Use GLTL to form closed-loop system matrix ATILDE

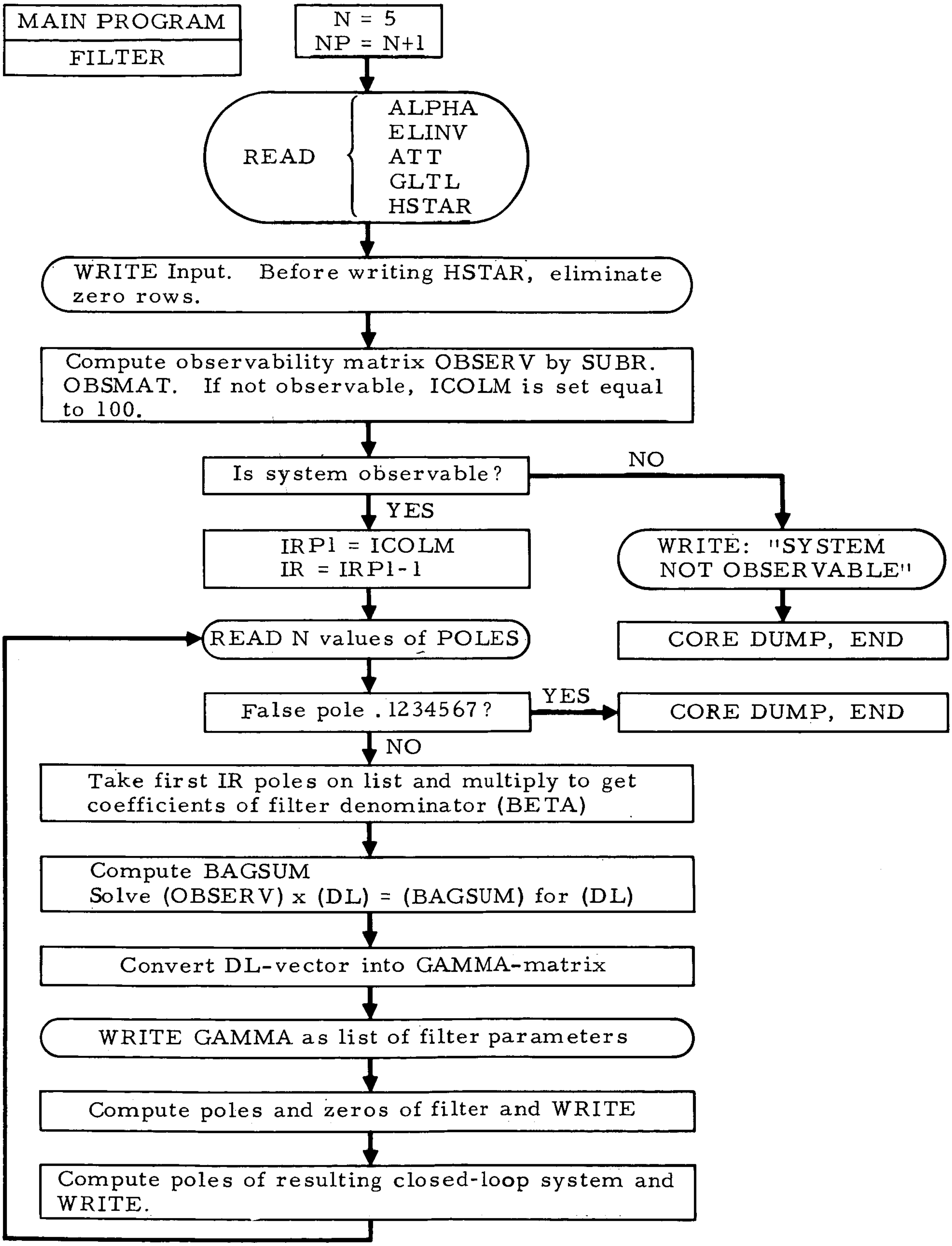
Find coefficients of closed-loop characteristic equation (ALPHA2) and compute resulting closed loop roots (CLROOT).

WRITE: ROOTS, GLTL, CLROOT

```
C OPTIMAL LITTLE-G MAIN PROGRAM (5TH ORDER)
C REQUIRES SPECIAL SUBROUTINES ALPHAS,MATMPY,MATPWR,ORDINV,POLYWG,SYNTH1
C
C*********************************************************************
      DIMENSION A(5,5),AAT(5,6),AB(5,1),AG(5,5),ALFOPT(12),ALTL(5,1),
     2          ALTLTR(1,5),ATP(5,5),ATR(5,5),ATILDE(5,5),ALPHA2(6),
     3          ALPHA(6),A2(5,5),APPROX(10),
     4          BLTL(5,1),C(5,5),C1(5,5),C2(5,5),CLROOT(5),COEFF(6),
     5          DSTAR(5,5),ELINV(5,5),EN(5,1),GLTL(5,1),S(5,5),
     6          GLTLTR(1,5),INDEX(6),OLROOT(5),OPTALF(6),ROOTS(10)
     7          ,TEMP(5)
C
C*********************************************************************
      COMPLEX  OLROOT,CLROOT,APPROX,ROOTS,ALFOPT
      EQUIVALENCE (AAT,BLTL),(AB,EN),(ATP,TEMP)
C*********************************************************************
C
      N=5
      NP=N+1
C
C*********************************************************************
C COMMENT ON INPUT
C         A- SYSTEM MATRIX
C         ALTL- ACTUATOR VECTOR
C         C1 AND C2 - PERFORMANCE MATRICES
C
      READ(1,1)MPRNTS
      READ(1,2)((A(I,J),J=1,N),I=1,N)
      READ(1,2)(ALTL(I,1),I=1,N)
      READ(1,2) ((C1(I,J),J=1,N),I=1,N), ((C2(I,J),J=1,N),I=1,N)
    1 FORMAT(10I2)
    2 FORMAT(5E10.0)
C
C*********************************************************************
C WRITE INPUT DATA
C
      WRITE(2,3) ((A(I,J),J=1,N),I=1,N)
    3 FORMAT(1H1,35X,28H CONTROL SYNTHESIS PROGRAM 2 //// 5X,9H A-MATRIX
     2 // (10X,5F15.4/)   )
      WRITE(2,4) (ALTL(I,1),I=1,N)
     2          ,((C1(I,J),J=1,N),I=1,N), ((C2(I,J),J=1,N),I=1,N)
    4 FORMAT(//5X,16H ACTUATOR VECTOR // 5(30X,F10.1//)///
     2 5X,21H PERFORMANCE MATRICES // 8X,14H DRIFT MINIMUM//
     3 5(10X,5E15.4//) / 8X,13H LOAD MINIMUM //5(10X,5E15.4//))
C
C*********************************************************************
C OBTAIN COEFFICIENTS OF OPEN LOOP CHARACTERISTIC EQUATION (ALPHAS) BY
C         CALLING SUBR ALPHAS.  IF MPRNTS IS 1, THE S-MATRICES WILL BE
C         WRITTEN BY THE SUBROUTINE.
C
      IF(MPRNTS.EQ.1)WRITE(2,14)
   14 FORMAT(1H1,24H S-MATRICES OF OPEN LOOP ////)
C
  701 CALL ALPHAS(N,A,ALTL,ALPHA,S,INDEX,MPRNTS,ELINV,ENORM)
C
      IF(MPRNTS.EQ.2)WRITE(2,15) ENORM
   15 FORMAT(1H1,30X,8H ENORM = E15.4,27H (SUM OF ABSOLUTE VALUES OF
     2        16H S-ZERO ELEMENTS ///)
```

```
C
C FIND ROOTS OF CHARACTERISTIC EQUATION (OPEN LOOP)
C
      DO 102 I=1,10
  102 APPROX(I)=(0.,0.)
      CALL ORDINV(NP,ALPHA)
      CALL ROOT1(N,ALPHA,OLROOT,APPROX,M)
      CALL ORDINV(NP,ALPHA)
C
C*********************************************************************
C WRITE RESULTS OF SUBR ALPHAS (ALPHA,OLROOT AND ELINV)
C
      WRITE(2,5)N,(ALPHA(I),I=1,NP),(OLROOT(I),I=1,N)
    5 FORMAT(/5X,34H OPEN LOOP CHARACTERISTIC EQUATION //
     2 10X,45H COEFFICIENTS OF ASCENDING POWERS OF S ( 0 TO,I2,2H )//
     3 /10X,6E18.5////5X,43H ROOTS OF OPEN LOOP CHARACTERISTIC EQUATION
     4 /// 30X,5H REAL,15X,10H IMAGINARY//(20X,2E20.5//) )
C
C
      WRITE(2,16)M
   16 FORMAT(30X,I2,20H SIGNIFICANT FIGURES ///)
C
      WRITE(2,10)
   10 FORMAT(1H1)
      WRITE(2,6) ((ELINV(I,J),J=1,N),I=1,N)
    6 FORMAT(31H ELINV (USED IN FILTER PROGRAM) ///5(10X,5E20.6//))
C
C*********************************************************************
C READ PERFORMANCE WEIGHTING FUNCTIONS
C
  931 READ(1,2) CAPPA1,CAPPA2
      IF(CAPPA1.EQ.1234567.) GO TO 50
      DO 933 I=1,N
      DO 933 J=1,N
  933 C(I,J)=CAPPA1*C1(I,J)+CAPPA2*C2(I,J)
      WRITE(2,10)
      WRITE(2,905) CAPPA1,CAPPA2, ((C(I,J),J=1,N),I=1,N)
  905 FORMAT(5X,30H PERFORMANCE WEIGHTING FACTORS //
     2 / 30X,30H DRIFT MINIMIZING (KAPPA-ROOF) E15.3 /
     3 / 30X,30H LOAD MINIMIZING (KAPPA-TILDE) E15.3 /
     4 ///31H WEIGHTED PERFORMANCE INDEX - C  // 5(10X,5E15.4//)  )
C
C*********************************************************************
C COMPUTE COEFFICIENTS OF OPTIMAL CLOSED LOOP CHARACTERISTIC EQUATION
C         (OPTALF).  S-MATRICES WRITTEN IF MPRNTS IS 1.
C
      CALL POLYWG(ALPHA,ELINV,C,COEFF,N,ROOTS,APPROX,ALFOPT)
C
      DO 203 IJ=1,NP
      I=N+2-IJ
  203 OPTALF(I)=REAL(ALFOPT(IJ))
      DO 107 I=1,N
  107 ALTLTR(1,I)=ALTL(I,1)
C
C*********************************************************************
C FROM A,ALPHA,OPTALF AND THE TRANSPOSE OF ALTL, SUBR SYNTH1 COMPUTES
C         THE CONTROL VECTOR GLTL.
C
```

```
      CALL SYNTH1(N,A,ALPHA,OPTALF,ALTLTR,ATR,ATP,AAT,EN,BLTL,AB,GLTL,
     2           TEMP,A2,DSTAR)
      WRITE(2,10)
      WRITE(2,7) ((DSTAR(I,J),J=1,N),I=1,N)
    7 FORMAT(///23H CONTROLLABILITY MATRIX /// 5(10X,5E20.6//) )
C
C***********************************************************************
C
      DO 109 I=1,N
  109 GLTLTR(1,I)=GLTL(I,1)
      CALL MATMPY(ALTL,N,GLTLTR,N,1,AG)
      DO 111 I=1,N
      DO 111 J=1,N
  111 ATILDE(I,J)=A(I,J)+AG(I,J)
C
C***********************************************************************
C USE GLTL TO FORM CLOSED-LOOP SYSTEM MATRIX ATILDE.  AGAIN USE SUBR
C        ALPHAS TO FIND ACTUAL CLOSED-LOOP ROOTS ACHIEVED BY CONTROL
C        VECTOR.
C
      IF(MPRNTS.EQ.1)WRITE(2,23)
   23 FORMAT(1H1,26H S-MATRICES OF CLOSED LOOP ////)
C
      CALL ALPHAS(N,ATILDE,ALTL,ALPHA2,S,INDEX,MPRNTS,ELINV,ENORM)
C
      IF(MPRNTS.EQ.2)WRITE(2,15)ENORM
      WRITE(2,10)
      WRITE(2,8) ((ATILDE(I,J),I=1,N),J=1,N)
    8 FORMAT(/// 47H A-TILDE-TRANSPOSE (ATT USED IN FILTER PROGRAM) /
     2         //5(10X,5E20.6//) )
C
C***********************************************************************
C
      CALL ORDINV(NP,ALPHA2)
      CALL ROOT1(N,ALPHA2,CLROOT,APPROX,M)
      WRITE(2,10)
      WRITE(2,13) (ROOTS(I),I=1,N), (GLTL(I,1),I=1,N),(CLROOT(I),I=1,N)
   13 FORMAT( //5X,26H OPTIMAL CLOSED-LOOP ROOTS /// 30X,5H REAL,10X,
     1 10H IMAGINARY //5(20X,2E20.5//) //
     2  5X,50H COMPUTED FEEDBACK CONTROL VECTOR (TERMS 1 THRU N) ///
     3 5(35X,E15.5//)// 5X,45H CLOSED LOOP ROOTS ACHIEVED BY CONTROL VEC
     4TOR /// 5(20X,2E20.5//)  )
C
C***********************************************************************
C
      GO TO 931
   50 WRITE(2,51)
   51 FORMAT(1H1)
      CALL DUMP
      STOP
      END
```


MAIN PROGRAM
FILTER
N = 5
NP = N+1
READ
ALPHA
ELINV
ATT
GLTL
HSTAR
WRITE Input. Before writing HSTAR, eliminate zero rows.
Compute observability matrix OBSERV by SUBR. OBSMAT. If not observable, ICOLM is set equal to 100.
Is system observable?
NO
WRITE: "SYSTEM NOT OBSERVABLE"
CORE DUMP, END
YES
IRP1 = ICOLM
IR = IRP1-1
READ N values of POLES
False pole .1234567?
YES
CORE DUMP, END
NO
Take first IR poles on list and multiply to get coefficients of filter denominator (BETA)
Compute BAGSUM
Solve (OBSERV) x (DL) = (BAGSUM) for (DL)
Convert DL-vector into GAMMA-matrix
WRITE GAMMA as list of filter parameters
Compute poles and zeros of filter and WRITE
Compute poles of resulting closed-loop system and WRITE.

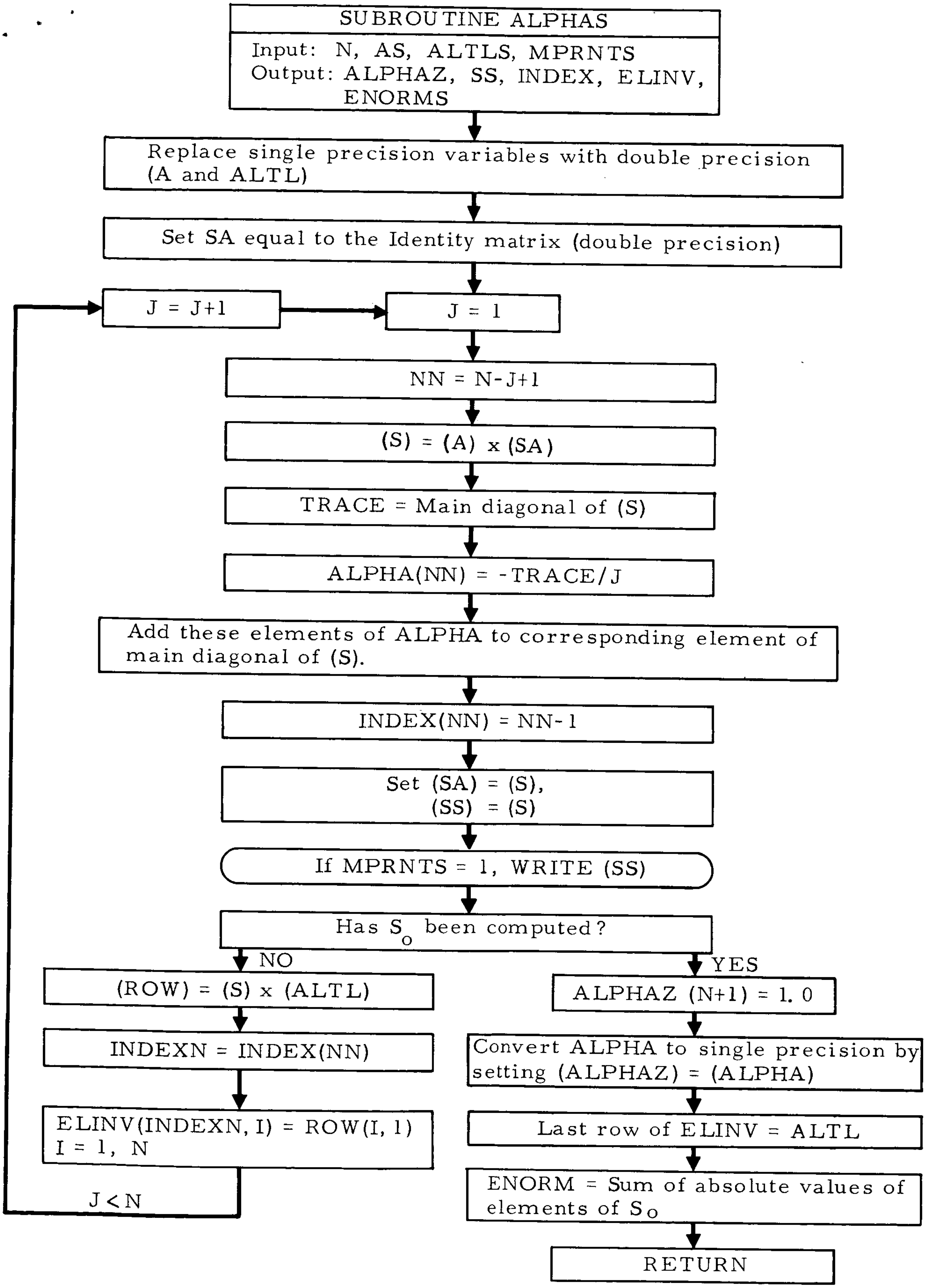
```
C  FILTER (5TH ORDER SYSTEM)
C  REQUIRES SUBROUTINES OBSMAT, MATMPY
C
C***********************************************************************
      DIMENSION A2(5,5),A3(5,5),A4(5,5),ALFSYS(11),ALPH(11),ALPHA(6),
     2 APPROX(5),ATGI(5,1),ATT(5,5),
     3 BAGSUM(5),BETA(6),BETA1(6), DL(5), ELINV(5,5), GAMMA(5,5),
     4 GAMMA1(5),GAMSA(11),GAMSAH(11),GLTL(5,1),  H(5,5),HPLA(11),
     A HSTAR(5,5),FILZRO(5) ,
     5 IJDLTL(5,5),NGLIST(5),NRROOT(10),OBSERV(5,5),POLE(5),POLES(5),
     6 SAH(5,5),SAH1(5),SUM(5), U(5,5),VK(5,6),VNXM(5,5),W(5)
C
      COMPLEX APPROX,BETA1,FILZRO,NRROOT,POLE,POLES
      EQUIVALENCE (H,HSTAR),(POLES,FILZRO)
C
C***********************************************************************
      N=5
      NP=N+1
C***********************************************************************
C COMMENT ON INPUT -
C        ALPHA- COEFFICIENTS OF OPEN-LOOP CHARACTERISTIC EQUATION
C        ELINV- AN OUTPUT OF A PROGRAM WHICH COMPUTED THE CONTROL VECTOR
C        ATT- TRANSPOSE OF THE A-TILDE MATRIX (FROM CNTRL1 OR CNTRL2)
C        GLTL- CONTROL VECTOR
C        HSTAR- ROWS(CARDS) REPRESENT SENSORS. THERE MUST BE N CARDS,
C                SOME BLANK IF LESS THAN N SENSORS.
C
      READ(1,1) (ALPHA(I),I=1,NP)
      READ(1,3) ((ELINV(I,J),J=1,N),I=1,N)
      READ(1,3) ((ATT(I,J),J=1,N),I=1,N)
      READ(1,3) (GLTL(I,1),I=1,N)
      READ(1,3) ((HSTAR(I,J),I=1,N),J=1,N)
    1 FORMAT(6E10.0)
    3 FORMAT(5E10.0)
C
C***********************************************************************
C WRITE INPUT
C
      WRITE(2,6 ) ((ATT(I,J),J=1,N),I=1,N)
    6 FORMAT(1H1,20X,37H DESIGN FILTER TO APPROXIMATE DESIRED
     2 26H SYSTEM POLE CONFIGURATION //// 25H A-TILDE-TRANSPOSE MATRIX
     3 ///(10X,5F18.6//) )
      WRITE(2,7) (GLTL(I,1),I=1,N)
    7 FORMAT(//26H G-LITTLE (CONTROL VECTOR) /// (30X,F20.6//) )
C
C DISCARD ZERO-ROWS OF HSTAR
      JJ=0
      DO 401 J=1,N
      HSUM=0.
      DO 102 I=1,N
  102 HSUM=HSUM + H(I,J)
      IF(HSUM.EQ.0.) GO TO 401
      JJ=JJ+1
      DO 103 I=1,N
  103 H(I,JJ)=H(I,J)
  401 CONTINUE
      M=JJ
C
      WRITE(2,10)
   10 FORMAT(1H1)
      WRITE(2,8)
    8 FORMAT(13H HSTAR MATRIX //)
```

```
      DO 51 I=1,M
       STO      SUBRLS+M,1
   18 FORMAT(/1 X,7HSENSOR I1,5X,7(F15.3) )
   51 WRITE(2,18) I,(HSTAR(J,I),J=1,N)
      WRITE(2,19) (ALPHA(I),I=1,NP)
   19 FORMAT(////50H COEFFICIENTS OF OPEN LOOP CHARACTERISTIC EQUATION
     2        26H (POWERS OF S FROM 0 TO N) ///20X,6E17.5)
      WRITE(2,9) ((ELINV(I,J),J=1,N),I=1,N)
    9 FORMAT(1H1,13H ELINV MATRIX /// 5(20X,5E18.7//)  )
C
C**********************************************************************
C COMPUTE OBSERVABILITY MATRIX (OBSERV). IF SYSTEM IS NOT OBSERVABLE,
C        ICOLM IS MADE EQUAL TO 100.  SUBR OBSMAT ALSO KEEPS TRACK OF
C        GAMMA POSITIONS VIA THE M BY ICOLM MATRIX IJDLTL.
C
      CALL OBSMAT(N,ATT,M,H,A2,A3,VNXM,SUM,U,W,NGLIST,IJDLTL,OBSERV,
     2            ICOLM)
      WRITE(2,11) ((OBSERV(I,J),J=1,N),I=1,N)
   11 FORMAT(1H1,21H OBSERVABILITY MATRIX /// (10X,5E20.6) )
      IF(ICOLM.EQ.100)GO TO 55
C
C ICOLM (IF NOT 100) DETERMINES NO. OF FILTER POLES REQUIRED
C
C**********************************************************************
      IRP1=ICOLM
      IR=IRP1-1
C
  402 CONTINUE
C
C**********************************************************************
C READ N CHOICES OF DESIRED FILTER POLES. PROGRAM WILL USE AS MANY AS
C        REQUIRED, STARTING FROM TOP OF LIST.
C
      READ(1,1) (POLES(I),I=1,N)
      POLERE=REAL(POLES(1))
C
C FALSE FILTER POLE MAY BE USED AS DATA TO CAUSE CORE DUMP.
      IF(POLERE.EQ.1234567.) GO TO 450
C
C**********************************************************************
C OBTAIN COEFFICIENTS OF FILTER DENOMINATOR (BETAS)
C
      DO 105 I=1,IR
  105 POLE(I)=POLES(I)
      CALL POLCO(IR,1.0,POLE,BETA1)
      DO 107 IJ=1,IRP1
      I=IRP1-IJ+1
  107 BETA(I)=REAL(BETA1(IJ))
C
C**********************************************************************
C PUT OBSERV AND BAGSUM IN FORM SUITABLE FOR SIMULTANEOUS EQUATION SUB-
C        ROUTINE (MATS).  SOLVE (OBSERV)X(DL)=(BAGSUM) FOR DL.
C
      DO 109 J=1,N
  109 BAGSUM(J)=0.
      DO 110 I=1,N
      DO 110 J=1,N
  110 A4(I,J)=0.
      DO 111 I=1,N
  111 A4(I,I)=1.
      DO 405 I=1,IRP1
      CALL MATMPY(A4,N,GLTL,1,N,ATGI)
```

```
      DO 112 J=1,N
  112 BAGSUM(J ) =BAGSUM(J)   +BETA(I)*ATGI(J,1)
      CALL MATMPY(ATT,N,A4,N,N,A3)
      DO 114 K=1,N
      DO 114 J=1,N
  114 A4(K,J)=A3(K,J)
  405 CONTINUE
      DO 113 I=1,N
      DO 113 J=1,N
  113 VK(I,J)= OBSERV(I,J)
      DO 115 I=1,N
  115 VK(I,N+1)= BAGSUM(I)
      CALL MATS(VK,DL,N,1)
C
C**********************************************************************
C THE IJDLTL MATRIX IS USED TO CONVERT THE VECTOR DL INTO A MATRIX GAMMA
C
      DO 117 J=1,ICOLM
      DO 117 I=1,M
  117 GAMMA(I,J)=0.
      JJ=1
      DO 406 J=1,ICOLM
      DO 406 I=1,M
      II=IJDLTL(I,J)+1
      GO TO (406,119),II
  119 GAMMA(I,J)=DL(JJ)
  305 JJ=JJ+1
  406 CONTINUE
C
C**********************************************************************
C WRITE THE PARAMETERS OF THE MULTIPORT FILTER AS COEFFICIENTS OF S.
C
      WRITE(2,12)M,IR
   12 FORMAT(1H1,50X,18H FILTER PARAMETERS //51X,I2,14H - PORT FILTER /
     2 /// 55H POLYNOMIAL COEFFICIENTS IN ASCENDING POWERS OF S (0 TO,
     3 I2,2H ) / )
      DO 407 I=1,M
      WRITE(2,13) I,(GAMMA(I,J),J=1,ICOLM)
   13 FORMAT(///5X,14H NUMERATOR NO.,I2 //5X,10E12.4)
  407 CONTINUE
   54 WRITE(2,14) (BETA(I),I=1,IRP1)
   14 FORMAT(///5X,19H COMMON DENOMINATOR//5X,10E12.4)
      WRITE(2,10)
C
C COMPUTE POLES AND ZEROS OF THE FILTER AND WRITE.
      WRITE(2,17)
   17 FORMAT(50X,13H FILTER ROOTS //   )
      DO 412 JJJ=1,M
      DO 141 I=1,N
  141 APPROX(I)=(0.,0.)
      JP=0
      DO 411 IJ=1,ICOLM
      I=ICOLM+1-IJ
      GAM1=GAMMA(JJJ,I)
      GAM2=ABS(GAM1)
      IF(GAM2.LE..00001.AND.JP.EQ.0) GO TO 411
      JP=JP+1
      GAMMA1(JP)=GAM1
  411 CONTINUE
      NA=JP-1
      IF(NA.LE.1)GO TO 412
      CALL ROOT1(NA,GAMMA1,FILZRO,APPROX,MSIG)
```

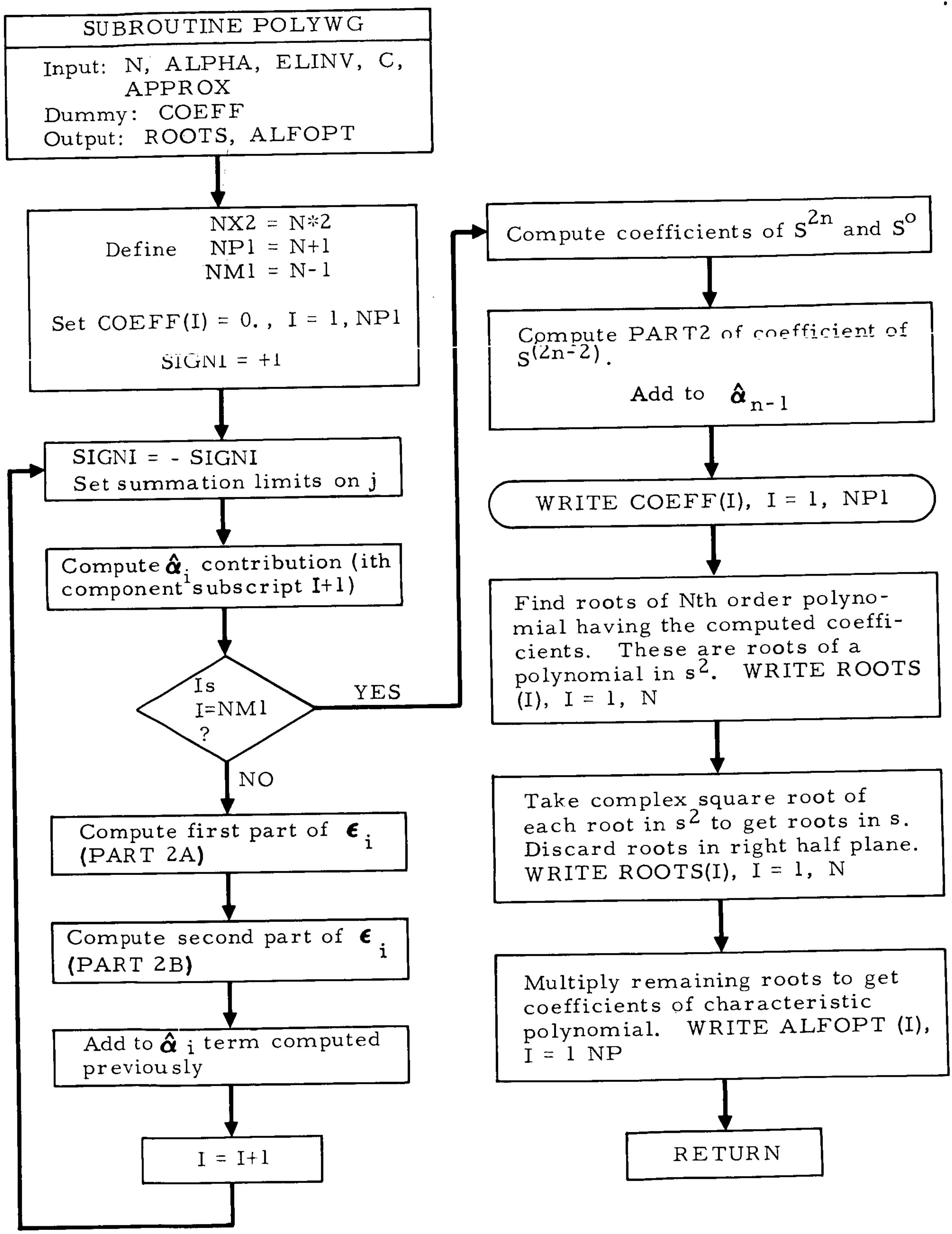
```
   21 FORMAT(///5X,14H NUMERATOR NO. I2/  (30X,2E20.6) )
      WRITE(2,21) JJJ,(FILZRO(J),J=1,NA)
  412 CONTINUE
   22 FORMAT(////5X,19H COMMON DENOMINATOR / (30X,2E20.6) )
      WRITE(2,22)(POLE(I),I=1,IR)
C
C***********************************************************************
C USING THE COMPUTED FILTER, DETERMINE THE POLES OF THE CLOSED LOOP.
C
      NSAH=N-1
      NGAM=ICOLM-1
      NGSP1=NSAH+NGAM+1
      CALL MATMPY(ELINV,N,H,M,N,SAH)
      DO 121 I=1,NGSP1
  121 GAMSAH(I)=0.
      DO 129 I=1,M
      DO 123 K=1,ICOLM
  123 GAMMA1(K)=GAMMA(I,K)
      DO 125 J=1,N
  125 SAH1(J)=SAH(J,I)
      DO 127 II=1,NGSP1
      GAMSA(II)=0.
      I1=II-NSAH
      JGAMIN = MAX0(I1,1)
      I2=NGAM+1
      JGAMAX= MIN0(I2,II)
      DO 127 JGAM=JGAMIN,JGAMAX
      JSAH=II+1-JGAM
  127 GAMSA(II)=GAMSA(II)+GAMMA1(JGAM)*SAH1(JSAH)
      DO 129 II=1,NGSP1
  129 GAMSAH(II)=GAMSAH(II)+GAMSA(II)
C
C***********************************************************************
      NPRP1=N+IR+1
      DO 133 I=1,NPRP1
      ALFSYS(I)=0.
      I1=I-N
      JMIN=MAX0(I1,1)
      I2=IR+1
      JMAX=MIN0(I2,I)
      DO 133 J=JMIN,JMAX
      K=I+1-J
  133 ALFSYS(I)=ALFSYS(I) + BETA(J)*ALPHA(K)
C
      NPR=N+IR
      NGSP2=NGSP1+1
      DO 135 I=NGSP2,NPRP1
  135 GAMSAH(I)=0.
      DO 137 I=1,NPRP1
  137 ALPH(I)=ALFSYS(I)-GAMSAH(I)
      DO 139 I=1,NPRP1
      J=NPRP1+1-I
  139 HPLA(I)=ALPH(J)
      CALL ROOT1(NPR,HPLA,NRROOT,APPROX,ISIG)
      WRITE(2,15) (NRROOT(I),I=1,NPR)
   15 FORMAT(1H1,20X,25H RESULTING POLE POSITIONS///(15X,2E20.6//)  )
C
C***********************************************************************
```

```
      GO TO 402
   55 WRITE(2,16)
   16 FORMAT(/////22H SYSTEM NOT OBSERVABLE)
  450 CONTINUE
      WRITE(2,10)
      CALL DUMP
      STOP
      END
```


SUBROUTINE ALPHAS
Input: N, AS, ALTLS, MPRNTS
Output: ALPHAZ, SS, INDEX, ELINV, ENORMS
Replace single precision variables with double precision (A and ALTL)
Set SA equal to the Identity matrix (double precision)
J = J+1
J = 1
NN = N-J+1
(S) = (A) x (SA)
TRACE = Main diagonal of (S)
ALPHA(NN) = -TRACE/J
Add these elements of ALPHA to corresponding element of main diagonal of (S).
INDEX(NN) = NN-1
Set (SA) = (S),
(SS) = (S)
If MPRNTS = 1, WRITE (SS)
Has S_o been computed?
NO
(ROW) = (S) x (ALTL)
INDEXN = INDEX(NN)
ELINV(INDEXN, I) = ROW(I, 1)
I = 1, N
J < N
YES
ALPHAZ (N+1) = 1. 0
Convert ALPHA to single precision by setting (ALPHAZ) = (ALPHA)
Last row of ELINV = ALTL
ENORM = Sum of absolute values of elements of S_o
RETURN

```
COMPUTE ALPHAS WITH DOUBLE PRECISION ARITHMETIC (INPUT,OUTPUT S.P.)
C REQUIRES SPECIAL SUBROUTINES MATMPY,DMATMP
C*********************************************************************
      SUBROUTINE ALPHAS(N,AS,ALTLS,ALPHAZ,SS,INDEX,MPRNTS,ELINV,ENORMS)
C
      DIMENSION AS(N,N),ALTLS(N,1),ALPHAZ(N),SS(N,N),INDEX(N),
     2 ELINV(N,N),
     3  A(5,5),ALTL(5,1),ALPHA(6),S(5,5),SA(5,5),ROW(5,1)
C
C COMMENT ON DIMENSIONING -  ALTHOUGH DIMENSIONED FOR N=5, THE WRITE
C         STATEMENTS ARE GENERAL FOR N LESS THAN 25.
C
      DOUBLE PRECISION A,ALTL,ALPHA,S,SA,TRACE,DFLOTJ,ROW,ENORM
C
C*********************************************************************
C IF MPRNTS=1, S-MATRICES ARE TO BE WRITTEN
      GO TO (1,2),MPRNTS
    1 IPRNT=1
      IF(N.GT.10) IPRNT=2
      IF(N.GT.20) IPRNT=3
    2 CONTINUE
C
C GET INPUT INTO DOUBLE PRECISION
      DO 101 I=1,N
      ALTL(I,1)=ALTLS(I,1)
      DO 101 J=1,N
  101 A(I,J)=AS(I,J)
      DO 3 I=1,N
      DO 3 J=1,N
    3 SA(I,J)=0.D0
      DO 4 K=1,N
    4 SA(K,K)=1.D0
C
C*********************************************************************
C USE LEVERRIER ALGORITHM TO COMPUTE SUCCESSIVE ALPHAS AND S-MATRICES
C         (N-1 THRU 0)
C
      DO 42 J=1,N
      NN=N-J+1
      CALL DMATMP(A,N,SA,N,N,S)
      TRACE=0.D0
      DO 6 K=1,N
    6 TRACE= TRACE + S(K,K)
      FLOTJ= FLOAT(J)
      DFLOTJ=FLOTJ
      ALPHA(NN) = -TRACE/DFLOTJ
      DO 8 K=1,N
    8 S(K,K)=S(K,K)+ ALPHA(NN)
      INDEX(NN)=NN-1
      DO 10 I=1,N
      DO 10 JJ=1,N
   10 SA(I,JJ)= S(I,JJ)
      DO 103 I=1,N
      DO 103 JD=1,N
  103 SS(I,JD)=S(I,JD)
C
C IF MPRNTS=1,WRITE S-MATRIX
      GO TO(51,405),MPRNTS
   51 WRITE(2,12) INDEX(NN)
   12 FORMAT(///5X,7H INDEX= I3//)
   22 DO 25 K=1,IPRNT
      JMIN=(K-1)*10 + 1
```

```
       MAX=K*10
       JMAX=MINO(N,MAX)
      WRITE(2,13) JMIN,JMAX
   13 FORMAT(5X,8H COLUMNS,I3,5H THRU,I3//)
      DO 25 I=1,N
   25 WRITE(2,14) (SS(I,JM),JM=JMIN,JMAX)
   14 FORMAT(10X,10E12.4/)
  405 CONTINUE
C
C THE ZERO-TH S-MATRIX SHOULD BE EQUAL TO ZERO. IT IS COMPUTED AS A
C        CHECK ON ROUNDOFF ERROR. AFTER IT IS COMPUTED, USE OF ALGORITHM
C        IS ENDED.
C
      IF(INDEX(NN))42,42,31
   31 CALL DMATMP(S,N,ALTL,1,N,ROW)
      INDEXN=INDEX(NN)
      DO 32 I=1,N
   32 ELINV(INDEXN,I)= ROW(I,1)
   42 CONTINUE
C
C************************************************************************
C GATHER LOOSE ENDS
C
      ALPHAZ(N+1)=1.0
      DO 105  I=1,N
  105 ALPHAZ(I)=ALPHA(I)
      INDEX(N+1)=N
      DO 43 I=1,N
   43 ELINV(N,I)= ALTL(I,1)
      ENORM=0.
      DO 44 I=1,N
      DO 44 J=1,N
   44 ENORM=ENORM+DABS(S(I,J))
      ENORMS=ENORM
      IF(MPRNTS.EQ.1)WRITE(2,100)
  100 FORMAT(1H1)
      RETURN
      END
```


SUBROUTINE POLYWG
Input: N, ALPHA, ELINV, C, APPROX
Dummy: COEFF
Output: ROOTS, ALFOPT
Define NX2 = N*2
NP1 = N+1
NM1 = N-1
Set COEFF(I) = 0., I = 1, NP1
SIGNI = +1
SIGNI = - SIGNI
Set summation limits on j
Compute $\hat{\alpha}_i$ contribution (ith component subscript I+1)
Is I=NM1 ?
YES
NO
Compute first part of ϵ_i (PART 2A)
Compute second part of ϵ_i (PART 2B)
Add to $\hat{\alpha}_i$ term computed previously
I = I+1
Compute coefficients of S^{2n} and S^o
Compute PART2 of coefficient of $S^{(2n-2)}$.
Add to $\hat{\alpha}_{n-1}$
WRITE COEFF(I), I = 1, NP1
Find roots of Nth order polynomial having the computed coefficients. These are roots of a polynomial in s^2. WRITE ROOTS (I), I = 1, N
Take complex square root of each root in s^2 to get roots in s. Discard roots in right half plane. WRITE ROOTS(I), I = 1, N
Multiply remaining roots to get coefficients of characteristic polynomial. WRITE ALFOPT (I), I = 1 NP
RETURN

```
      SUBROUTINE POLYWG(ALPHA,ELINV,C,COEFF,N,ROOTS,APPROX,ALFOPT)
C
C REQUIRES SPECIAL SUBROUTINE ORDINV
C
C***********************************************************************
C THIS SUBROUTINE IS A MECHANIZATION OF EQUATIONS (44) AND (57) OF THE
C        PAPER ON HIGHER ORDER SYSTEM DESIGN BY BASS AND GURA.
C***********************************************************************
C
      DIMENSION ALPHA(1),COEFF(1),ELINV(N,N),C(N,N),ROOTS(2),ALFOPT(1)
C
C NOTE ON DIMENSIONS -(1) INDICATES CALLING ROUTINE TO BE DIMENSIONED
C        (N+1).   (2) INDICATES (N*2).
C
      COMPLEX   ROOTS,APPROX,ALFOPT
C
C***********************************************************************
      NX2=N*2
      NP1=N+1
      DO 101 I=1,NP1
  101 COEFF(I)=0.
      SIGNI=1.
      NM1=N-1
C
C***********************************************************************
C
C COMPUTE EVEN COEFFICIENTS OF THE POLYNOMIAL DELTA(2N).  ODD COEFFIC-
C        IENTS ARE ZERO.
C
      DO 430  I =1,NM1
      SIGNI= -SIGNI
      JMIN=1
      JMAX= I
      IX2 = I*2
      IF( IX2.GT.N ) JMIN= IX2-N+1
C
C COMPUTE ALPHA-HAT TERM (PART 1)
C
      SIGNJ=+1.
      JEVN=JMIN-2*(JMIN/2)
      IF(JEVN.GT.0) SIGNJ=-1.
      DO 105 JP1=JMIN,JMAX
      SIGNJ=-SIGNJ
      JS=IX2-JP1+2
  105 COEFF(I+1)=COEFF(I+1)+ALPHA(JP1)*ALPHA(JS)*SIGNJ
      COEFF(I+1)=2.*COEFF(I+1)+SIGNI*ALPHA(I+1)*ALPHA(I+1)
C***********************************************************************
      IF(I.EQ.NM1) GO TO 430
      JMIN=1
      IF(IX2.GT.NM1) JMIN=IX2-N+2
      JEVN=JMIN-2*(JMIN/2)
      SIGNJ=1.
      IF(JEVN.GT.0) SIGNJ=-1.
      PART2A=0.
C
C COMPUTE EPSILON TERM ( 2 PARTS)
      DO 110 JP1=JMIN,JMAX
      SIGNJ=-SIGNJ
      ANS=0.
      JS=IX2-JP1+2
      DO 107 JJ=1,N
```

```
      DO 107 KK=1,N
  107 ANS=ANS+ELINV(JP1,JJ)*C(JJ,KK)*ELINV(JS,KK)
  110 PART2A=PART2A + SIGNJ*ANS
      PART2B=0.
      DO 120 JJ=1,N
      DO 120 KK=1,N
  120 PART2B=PART2B + ELINV(I+1,JJ)*C(JJ,KK)*ELINV(I+1,KK)
C
C ADD EPSILON TERM TO ALPHA-HAT TERM
      COEFF(I+1)=COEFF(I+1) + 2.*PART2A + PART2B*SIGNI
  430 CONTINUE
C
C
C***********************************************************************
C
C 2N-TH AND ZERO-TH TERM ARE COMPUTED SEPARATELY
C
C
C COEFFICIENT OF S**2N -
      NEVN=N-2*(N/2)
      COEFF(N+1)=-1.
      IF(NODD.GT.0)COEFF(N+1)=1.
      COEFF(1)=0.
      DO 124 JJ=1,N
      DO 124 KK=1,N
  124 COEFF(1)= COEFF(1) + ELINV(1,JJ)*C(JJ,KK)*ELINV(1,KK)
      COEFF(1)= COEFF(1) + ALPHA(1)*ALPHA(1)
C
C***********************************************************************
C PART 2 OF COEFFICIENT OF S**(2N-2)
C
      SIGNJ=-COEFF(N+1)
C     PART2=0.
      DO 126 JJ=1,N
      DO 126 KK=1,N
  126 PART2 = ELINV(N,JJ)*C(JJ,KK)*ELINV(N,KK)
C
C ADD PART2 TO PART1 ((N-1) TERM)
      COEFF(N)=COEFF(N)+SIGNJ*PART2
C
C***********************************************************************
C WRITE HEADING FOR INTERNAL POLYWG WRITE STATEMENTS.
      WRITE(2,11)
   11 FORMAT(1H1 15X,46H SUBROUTINE POLYWG PRINTS COEFFICIENTS OF EVEN
     2         35H POWERS OF S (ODD POWERS ARE ZERO), /
     3         20X,50H N ROOTS OF POLYNOMIAL FORMED BY EVEN COEFS. ONLY,
     4         23H N L.H.P. ROOTS OF 2NTH / 20X,18H ORDER POLYNOMIAL,
     5         52H AND THE COEFFICIENTS GENERATED BY THE L.H.P. ROOTS.
     6         )
C
C WRITE COEFFICIENTS COMPUTED ABOVE.
C
      WRITE(2,12)
   12 FORMAT(///// 20X,31H POWER OF S        COEFFICIENT // )
      DO 51 I=1,NP1
      NNX2=2*I-2
   51 WRITE(2,13)NNX2,COEFF(I)
   13 FORMAT(24X,I3,12X,E15.8)
C
C***********************************************************************
C
C FIND ROOTS OF POLYNOMIAL HAVING THE N+1 COEFFICIENTS PREVIOUSLY COM-
```

```
C          PUTED.  THESE ARE THE ROOTS OF THE FORM (S**2 - ROOT).
C
      CALL ORDINV(NP1,COEFF)
      CALL ROOT1(N,COEFF,ROOTS,APPROX,M)
      WRITE(2,14) (ROOTS(I),I=1,N)
   14 FORMAT(////29H ROOTS FROM EVEN COEFFICIENTS /
     2        //(20X,2E20.7) )
C
C***********************************************************************
C TAKE THE COMPLEX SQUARE ROOTS (OF THE PREVIOUS ROOTS) WHICH ARE IN THE
C          LEFT HALF PLANE.  THESE ARE THE OPTIMAL CLOSED-LOOP ROOTS.
C
      DO 404 I=1,N
      ROOTS(I)=CSQRT(ROOTS(I))
      ROOTRE=REAL(ROOTS(I))
      IF(ROOTRE.LT.0.)GO TO 404
      ROOTS(I)=-ROOTS(I)
  404 CONTINUE
      WRITE(2,10)
   10 FORMAT(1H1)
      WRITE(2,15) (ROOTS(I),I=1,N)
   15 FORMAT(10X,30HROOTS WITH NEGATIVE REAL PARTS // (2E20.7))
C
C***********************************************************************
C MULTIPLY THE ROOTS TO GET THE COEFFICIENTS OF THE DESIRED CLOSED-LOOP
C          CHARACTERISTIC POLYNOMIAL.
C
      CALL POLCO(N,1.0,ROOTS,ALFOPT)
      WRITE(2,16)
   16 FORMAT(///50H COEFFICIENTS OF POLYNOMIAL GENERATED BY ROOTS IN ,
     1        15HLEFT HALF PLANE //10X,10HPOWER OF S 26X,11HCOEFFICIENT /
     2        40X,4HREAL 14X,9HIMAGINARY // )
      NP=N+1
      DO 53 I=1,NP
      NMI=NP-I
   53 WRITE(2,17)NMI,ALFOPT(I)
   17 FORMAT(14X,I3,12X,2E20.8)
C
      RETURN
      END
```

SUBROUTINE SYNTH1

Input: N, A, ALPHA, DSIRD, ALTLTR
Dummy: ATR, ATP, AAT, EN, BLTL, AB, TEMP, A2
Output: GLTL, DSTAR

Compute DSTAR by
$D^* = (a, Aa, \ldots, A^{n-1}a)$

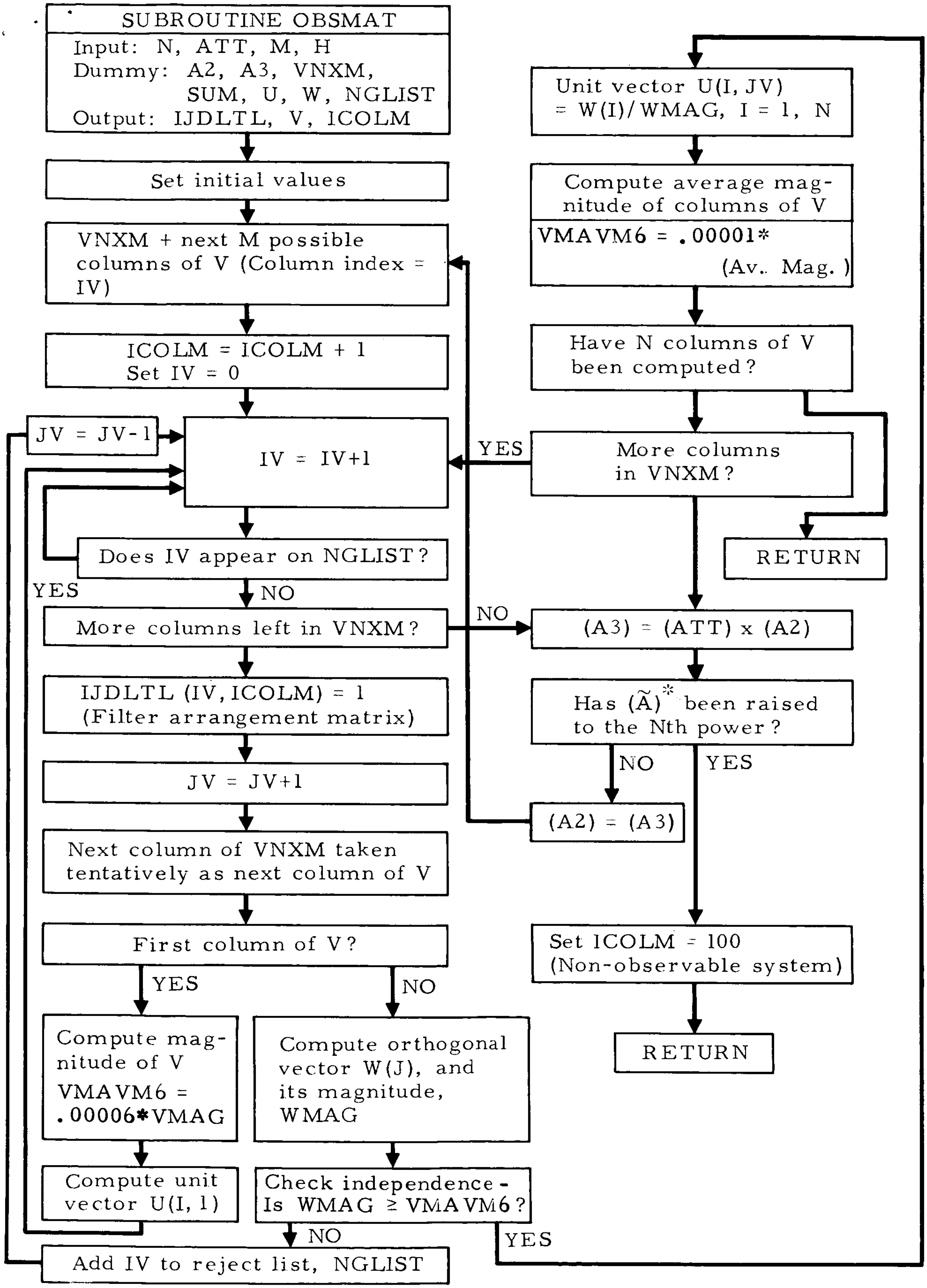
Compute BLTL (b) by solving
$D^*b = e^n$
where $e^n = (0, 0, 0, \ldots 0, 1)^*$

Compute GLTL by
$$g = -\sum_{i=1}^{n} (\tilde{\alpha}_{i-1} - \alpha_{i-1})(A^*)^{i-1} b$$

RETURN

```
      SUBROUTINE SYNTH1(N,A,ALPHA,DSIRD,ALTLTR,ATR,ATP,AAT,EN,BLTL,AB,
     2                    GLTL,TEMP,A2,DSTAR)
C
C     EQUIVALENCE (AAT,BLTL),(EN,AB),(ATP,TEMP)
C
      DIMENSION A(N,N),ALPHA(2),ALTLTR(1,N),AAT(N,N),ATP(N,N),ATR(N,N),
     2          AB(N,1),A2(N,N),BLTL(N,1),DSTAR(N,N),DSIRD(2),EN(N,1),
     3          GLTL(N,1),TEMP(N)
C
C COMMENT ON DIMENSION - (2) INDICATES DIMENSION OF N+1 IN MAIN PROGRAM.
C
C***********************************************************************
      DO 1 I=1,N
      DO 1 J=1,N
      ATP(I,J)=0.
    1 ATR(I,J)=A(J,I)
      DO 2 I=1,N
    2 ATP(I,I)=1.
C
C***********************************************************************
C COMPUTE DSTAR (TRANSPOSE OF CONTROLLABILITY MATRIX).  AAT IS SET EQUAL
C        TO DSTAR TO AVOID DESTROYING DSTAR IN THE SIMULTANEOUS EQUATION
C        SUBROUTINE (SIMEQ).
C
      DO 7 K=1,N
      DO 101 I=1,N
      DSTAR(K,I)=0.
      DO 101 J=1,N
  101 DSTAR(K,I)=DSTAR(K,I)+ALTLTR(1,J)*ATP(J,I)
      DO 3 I=1,N
    3 AAT(K,I)=DSTAR(K,I)
      CALL MATMPY(ATP,N,ATR,N,N,A2)
      DO 4 I=1,N
      DO 4 J=1,N
    4 ATP(I,J)=A2(I,J)
    7 CONTINUE
C
C***********************************************************************
C SOLVE  (DSTAR)X(BLTL)=(EN) FOR BLTL
C
      DO 8 I=1,N
    8 EN(I,1)=0.
      EN(N,1)=1.
      SCALE=1.0
      CALL SIMEQ(N,N,1,AAT,EN,SCALE,TEMP,MM)
C
C THE SOLUTION VECTOR BLTL IS THE FIRST COLUMN OF AAT (SEE COMMENT ON
C        EQUIVALENCE STATEMENT)
C
      GO TO (402,52,53),MM
   52 WRITE(2,14)
   14 FORMAT(////42H UNDERFLOW OR OVERFLOW IN SUBROUTINE SIMEQ)
      GO TO 402
   53 WRITE(2,15)
   15 FORMAT(////51H SYSTEM NOT CONTROLLABLE (DSTAR-MATRIX IS SINGULAR))
  402 CONTINUE
C
C***********************************************************************
C COMPUTE THE CONTROL VECTOR GLTL.  THIS IS THE PRIMARY RESULT OF THIS
C        SUBROUTINE AND OF THE MAIN PROGRAM.
C
      DO  9 I=1,N
```

```
    9 GLTL(I,1)=0.
      DO 10 I=1,N
      DO 10 J=1,N
   10 ATP(I,J)=0.
      DO 11 I=1,N
   11 ATP(I,I)=1.
      DO 13 K=1,N
      CALL MATMPY(ATP,N,BLTL,1,N,AB)
      DO 12 I=1,N
   12 GLTL(I,1)=GLTL(I,1) + (ALPHA(K)-DSIRD(K))*AB(I,1)
      CALL MATMPY(ATP,N,ATR,N,N,A2)
      DO 13 I=1,N
      DO 13 J=1,N
   13 ATP(I,J)=A2(I,J)
C
      RETURN
      END
```


SUBROUTINE OBSMAT
Input: N, ATT, M, H
Dummy: A2, A3, VNXM, SUM, U, W, NGLIST
Output: IJDLTL, V, ICOLM
Set initial values
VNXM + next M possible columns of V (Column index = IV)
ICOLM = ICOLM + 1
Set IV = 0
JV = JV-1
IV = IV+1
Does IV appear on NGLIST?
YES
NO
More columns left in VNXM?
NO
IJDLTL (IV, ICOLM) = 1
(Filter arrangement matrix)
JV = JV+1
Next column of VNXM taken tentatively as next column of V
First column of V?
YES
NO
Compute magnitude of V
VMAVM6 = .00006*VMAG
Compute orthogonal vector W(J), and its magnitude, WMAG
Compute unit vector U(I, 1)
Check independence - Is WMAG ≥ VMAVM6?
NO
YES
Add IV to reject list, NGLIST
Unit vector U(I, JV) = W(I)/WMAG, I = 1, N
Compute average magnitude of columns of V
VMAVM6 = .00001*
(Av. Mag.)
Have N columns of V been computed?
More columns in VNXM?
YES
RETURN
(A3) = (ATT) x (A2)
Has (Ã)* been raised to the Nth power?
NO
YES
(A2) = (A3)
Set ICOLM = 100
(Non-observable system)
RETURN

```
C     OBSERVABILITY MATRIX
      SUBROUTINE OBSMAT(N,ATT,M,H,A2,A3,VNXM,SUM,U,W,NGLIST,IJDLTL,
     2                  V,ICOLM)
      DIMENSION ATT(N,N),H(N,N),A2(N,N),A3(N,N),VNXM(N,N),U(N,N),
     2          IJDLTL(N,N),V(N,N),SUM(N),W(N),NGLIST(N)
C***********************************************************************
C
      NMM=N-M
      DO 4 I=1,M
      DO 4 J=1,N
    4 IJDLTL(I,J)= 0
C
C***********************************************************************
C     TAKE FIRST M COLUMNS (NXM H-MATRIX)
C
      DO 5 I=1,N
      DO 5 J=1,N
    5 A2(I,J)=0.
      DO 7 I=1,N
    7 A2(I,I)=1.
      JV=0
      IATP=1
      VMAGSM=0.
      VMAVM6=0.
      NGLIST(1)=0
      MLIST=1
      ICOLM=0
C
C***********************************************************************
    9 CALL MATMPY(A2,N,H,M,N,VNXM)
      ICOLM=ICOLM+1
      IV=0
C
C***********************************************************************
C     COMPARE COLUMN INDEX WITH N.G.LIST, ELIMINATE COLUMNS ON LIST.
C
   10 IV=IV+1
      DO 12 I=1,MLIST
      IF(NGLIST(I)-IV)12,10,12
   12 CONTINUE
      IF(IV-M-1)14,28,35
   14 CONTINUE
      IJDLTL(IV,ICOLM)=1
C
C***********************************************************************
C     TAKE A COLUMN AT A TIME
C
      JV=JV+1
      DO 15 I=1,N
   15 V(I,JV)=VNXM(I,IV)
      IF(JV-1)35,110,16
  110 VMAGSQ=0
      DO 111 I=1,N
  111 VMAGSQ=VMAGSQ+V(I,JV)*V(I,JV)
      VMAG=SQRT(VMAGSQ)
      VMAVM6 = .000006*VMAG
      DO 112 I=1,N
  112 U(I,JV)=V(I,JV)/VMAG
      GO TO 10
C
C***********************************************************************
C     COMPUTE ORTHOGONAL VECTOR W
```

```
C
   16 JVM1=JV-1
      DO 19 J=1,N
      SUM(J)=0.
      DO 18 I=1,JVM1
      DO 18 K=1,N
   18 SUM(J)=SUM(J)+U(K,I)*V(K,JV)*U(J,I)
   19 W(J)=V(J,JV)-SUM(J)
C
C*********************************************************************
C     IF MAGNITUDE OF W IS ZERO,REJECT COLUMN AND TRY NEXT
C
      WMAGSQ =C.
      DO 21 I=1,N
   21 WMAGSQ  =  WMAGSQ + W(I)**2
      WMAG    = SQRT(WMAGSQ)
      IF(WMAG  -  VMAVM6) 23,24,24
   23 MLIST=MLIST +1
      NGLIST(MLIST)=IV
      IJDLTL(IV,ICOLM)=0
      JV=JV-1
      GO TO 10
   24 CONTINUE
      DO 25 I=1,N
   25 U(I,JV)= W(I)/WMAG
      VMAGSQ=0.
      DO 26 I=1,N
   26 VMAGSQ = VMAGSQ + V(I,JV)*V(I,JV)
      VMAG=SQRT(VMAGSQ)
      VMAGSM= VMAGSM+VMAG
      VMAVM6= .000001*VMAGSM/FLOAT(JV)
      IF(JV-N)27,40,35
   27 IF(IV-M)10,28,28
C
C*********************************************************************
C     COMPUTE NEXT POWER OF A-TILDE-TRANSPOSE
C
   28 CALL MATMPY(ATT,N,A2,N,N,A3)
      IATP=IATP+1
      IF(IATP-N-1)29,32,35
   29 DO 30 I=1,N
      DO 30 J=1,N
   30 A2(I,J)=A3(I,J)
      GO TO 9
   32 ICOLM=100
   34 GO TO 40
   35 WRITE(2,36)
   36 FORMAT(/////27H ERROR IN OBSMAT SUBROUTINE)
   40 RETURN
      END
```

```
      SUBROUTINE DMATMP(A,NR,B,NC,N,C)
      DIMENSION A(NR,N),B(N,NC),C(NR,NC)
      DOUBLE PRECISION A,B,C
      DO 1 I=1,NR
      DO 1 K=1,NC
      C(I,K)=0.D0
      DO 1 J=1,N
    1 C(I,K)=C(I,K)+A(I,J)*B(J,K)
      RETURN
      END
```

```
      SUBROUTINE MATMPY(A,NR,B,NC,N,C)
      DIMENSION A(NR,N),B(N,NC),C(NR,NC)
      DO 1 I=1,NR
      DO 1 K=1,NC
      C(I,K)=0.0
      DO 1 J=1,N
    1 C(I,K)=C(I,K)+A(I,J)*B(J,K)
      RETURN
      END
```

```
      SUBROUTINE ORDINV(N,V)
      DIMENSION V(N)
      NHALF=N/2
      DO 1 I=1,NHALF
      N1=N+1-I
      A=V(N1)
      V(N1)=V(I)
    1 V(I)=A
      RETURN
      END
```

SYMBOL DICTIONARY

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
A(I, J)†	A	nxn matrix which defines the autonomous set of linear differential first order equations of the form $\dot{x} = Ax$, where x is an n-vector	CNTRL2 SYNTH1 ALPHAS
AAT(I, J)		A dummy matrix which is set equal to DSTAR to make use of the simultaneous equation subroutine SIMEQ in solving $D^*b = e^n$ for b (BLTL)	SYNTH1
AB(I, 1)	$(A^*)^i b$	Used in the computation of g.	SYNTH1
AG(I, J)	ag^*	Used to find $\tilde{A}$ by $\tilde{A} = A+ag^*$	CNTRL2

†Except as otherwise noted, subscripts run from 1 to N

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
ALFOPT(I)		Coefficients of optimal closed-loop characteristic equation as computed by POLYWG. They are complex but the imaginary part is non-zero only due to computing error. They are also in the wrong order. (See OPTALF)	POLYWG CNTRL2
ALFSYS(I) I=1, N+R+1		Coefficients of $\Delta(S)\Delta_{n-\nu}(S)$ in Eq. (28b)	FILTER
ALPH(I) I=1, N+R+1		Coefficients of $\tilde{\Delta}_{2n-\nu}$ in Eq. (28b)	FILTER
ALPHA(I) I=1, N+1	α_i i=o, n	Coefficients of the characteristic equation of the plant, $\Delta(S)$, in ascending powers of S from 0 (I=1) to n (I=N+1).	CNTRL2 FILTER SYNTH1 ALPHAS
ALPHAZ(I) I=1, N+1		Dummy variable used in ALPHAS to indicate single precision (ALPHA is double precision in ALPHAS)	ALPHAS

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
ALPHA2(I) I=1, N+1	$\tilde{\alpha}_i$ i=o, n	Coefficients of the closed loop characteristic equation $\tilde{\Delta}(S)$.	CNTRL2
ALTL(I, 1)	a	n-vector which is called the actuator vector. From the equation $\dot{x} = Ax + a\psi$	CNTRL2
ALTLS(I, 1)		Dummy used in ALPHAS to indicate single precision.	ALPHAS
ALTLTR(1, I)	a*	Transpose of ALTL	CNTRL2 SYNTH1
APPROX(I)		Guess at roots to assist ROOT1 subroutine (library).	CNTRL2 FILTER POLYWG
ATGI(I, 1)	$(A^*)^{j-1} g$	Intermediate variable used to compute d (DL)	FILTER
ATILDE(I, J)	$\tilde{A}$	Analogous to A, except it defines the closed loop system.	CNTRL2
ATP(I, J)		Dummy matrix used to represent successively higher powers of A^*	SYNTH1

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
ATR(I, J)	A^*	Transpose of A	SYNTH1
ATT(I, J)	$(\tilde{A})^*$	Transpose of the closed-loop system matrix	CNTRL2 FILTER
A2(I, J) A3(I, J) etc.		Dummy matrices	
BAGSUM(I)	$\Delta_{n-\nu}(\tilde{A})^* g$ $=Qr$	Intermediate variable used to compute d	FILTER
BETA(I) I=1, R+1	γ_i	Coefficients of S in common denominator of filter ($\Delta_{n-\nu}$).	FILTER
BLTL(I, 1)	b	Vector resulting from the solution of $D^* b = e^n$	CNTRL2 SYNTH1
c(I, J)	c	The weighted performance index.	CNTRL2 POLYWG
CLROOT(I)		Achieved closed loop poles	CNTRL2
C1(I, J), C2(I, J)		The performance indices of $c = K_1 C1 + K_2 C2$	CNTRL2

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
COEFF(I) I=1, N+1		Dummy array used to compute the optimal coefficients ALFOPT	POLYWG
DL(I)	d	The coefficients of the filter numerators arranged consecutively. See definition following Eq. (31).	FILTER
DSIRD(I) I=1, N+1	$\tilde{\alpha}_i$ $i=o, n$	Coefficients of the desired closed loop characteristic equation. Corresponds to OPTALF(I).	SYNTH1
DSTAR(I, J)	D^*	The transpose of the controllability matrix, D	CNTRL2 SYNTH1
ELINV(I, J)	L^{-1}	Inverse of L. $L^{-1} = (S_1 a, S_2 a, \ldots, S_n a)^*$	ALPHAS CNTRL2 FILTER
EN(I, 1)	e^n	$e^n = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$ -nth element	SYNTH1

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
ENORM		Error norm-sum of the absolute values of the elements of the final S-matrix computed (S_o) by ALPHAS. Used as a check on computational accuracy.	ALPHAS
FILZRO(I)		Used repeatedly to compute and write the filter zeros.	FILTER
GAMMA(I, J) I=1, M J=1, ICOLM	γ_{ij}	Each of M rows of GAMMA represents the coefficients of increasing powers of s of each of M filter numerators. I=1, 2, . . . , ICOLM corresponding to $s^o, s^1, \ldots s^r$	FILTER OBSMAT
GAMSAH(I) I=1, N+ICOLM+1		Coefficients of s in 2nd term of Eq. (28b), $\sum_{i=1}^{M} \Delta_{(i)}(s) \cdot \sum_{j=1}^{n} (h^i \cdot S_j a)\, s^{j-1}$	FILTER

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
GLTL(I, 1)	g	Control vector, from the equation $\dot{x} = Ax + a\psi$, where $\psi = g \cdot x$	CNTRL2 FILTER SYNTH1
GLTLTR(1, I)	g^*	Transpose of the control vector	CNTRL2
H(I, J)	H	The sensor matrix. Each of M sensors represented by a column of H.	FILTER OBSMAT
HPLA(I) I=1, N+Ṙ+1		The elements of ALPH(I) in reverse order	FILTER
HSTAR(I, J)	H^*	Transpose of H	FILTER
ICOLM		Number of columns in GAMMA(I, J). Tells maximum number of zeros of filter (ICOLM-1). Used as an indicator of non-observability by setting equal to 100.	FILTER OBSMAT
IJDLTL(I, J) I=1, M J=1, ICOLM		Filter arrangement matrix. Elements are either 1 or 0.	OBSMAT FILTER

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
INDEX(I) I=1, N+1		Used to allow print-out of subscript zero, which is not allowed in Fortran.	ALPHAS CNTRL2
IPRNT		Determines printing format in ALPHAS. Value depends on size of array to be printed.	ALPHAS
MPRNTS		Option to print S-matrices of Subr. Alphas decides by input value of MPRNTS. Printing occurs if MPRNTS = 1, does not if MPRNTS = 2.	ALPHAS
N	n	Order of the plant characteristic equation.	All
NGLIST(I) I=1, MLIST		List of rejected column numbers. Used to compute observability matrix.	OBSMAT
NRROOT(I) I=1, N+R		Closed loop poles of system with filter.	FILTER

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
OBSERV(I, J)	K	Observability matrix computed by OBSMAT. Dummy matrix in OBSMAT corresponding to OBSERV is V.	FILTER
OLROOT(I)		Open loop poles	CNTRL2
OPTALF(I) I=1, N+1		The real parts of coefficients ALFOPT, with order corrected.	CNTRL2
POLE(I) I=1, IR		The common filter poles. R poles taken from an arbitrary list of N provided as input.	FILTER
POLES(I)		The N poles available as common filter poles.	FILTER
ROW(I, 1)		Dummy array used in ALPHAS. ROW is a column matrix, but is a row in ELINV.	ALPHAS
S(I, J)	S	The numerator transfer matrix. Successive values of the matrix are computed as a part of computing ALPHA(I).	CNTRL2 ALPHAS

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
SA(I, J)		Dummy variable used in ALPHAS to replace S.	ALPHAS
SAH(J, I) J=1, N I=1, M	$h^i \cdot S_{ja}$	Coefficients of open loop numerators	FILTER
SCALE		A scale factor used in the library subroutine SIMEQ.	SYNTH1
SUM(I)		Dummy array used by OBSMAT in determination of column independence.	OBSMAT
TEMP(I)		Temporary storage required by library subroutine SIMEQ.	SYNTH1
TRACE	tr	Sum of the elements on the main diagonal of a matrix.	ALPHAS
U(I, J)		The columns are a series of unit vectors formed successively by OBSMAT in determination of column independence.	OBSMAT

Fortran Symbol	Mathematical Equivalent, if any	Definition	Used in
VK(I, J) I=1, N J=1, N+1		A dummy matrix used to represent OBSERV and BAGSUM in the library subroutine MATS.	FILTER
VNXM(I, J) I=1, N J=1, M	$(A^*)^i H$	Dummy matrix used successively in OBSMAT.	OBSMAT
W(I)		The orthogonal vector computed by OBSMAT. Recomputed successively.	OBSMAT

A-MATRIX

-0.0322	-0.0194	1.0000	-0.0211	-0.
-0.	-0.	1.0000	-0.	-0.
0.0693	-0.	-0.	-0.4740	-0.
-0.	-0.	-0.	-0.	1.0000
0.7620	-0.	-0.	-1760.5000	-3.3600

ACTUATOR VECTOR (a)

-0.
-0.
-0.
-0.
17.5

PERFORMANCE MATRICES

DRIFT MINIMUM ($\hat{C}$ or C1)

0.1040E-02	0.6240E-03	-0.	0.6800E-03	-0.
0.6240E-03	0.3750E-03	-0.	0.4100E-03	-0.
-0.	-0.	-0.	-0.	-0.
0.6800E-03	0.4100E-03	-0.	0.4450E-03	-0.
-0.	-0.	-0.	-0.	-0.

LOAD MINIMUM ($\tilde{C}$ or C2)

0.1610E 01	-0.	-0.	0.5600E 01	-0.
-0.	-0.	-0.	-0.	-0.
-0.	-0.	-0.	-0.	-0.
0.5600E 01	-0.	-0.	0.1610E 01	-0.
-0.	-0.	-0.	-0.	-0.

S-MATRICES OF OPEN LOOP

An option in Subroutine ALPHAS can print each S-matrix as it is computed. The matrix in storage is then replaced with the next matrix, etc.

INDEX= 4 (S_4)

COLUMNS 1 THRU 5

0.3360E 01	-0.1940E-01	0.1000E 01	-0.2110E-01	-0.
-0.	0.3392E 01	0.1000E 01	-0.	-0.
0.6930E-01	-0.	0.3392E 01	-0.4740E-00	-0.
0.	0.	0.	0.3392E 01	0.1000E 01
0.7620E 00	-0.	-0.	-0.1760E 04	0.3220E-01

INDEX= 3 (S_3)

COLUMNS 1 THRU 5

0.1760E 04	-0.6518E-01	0.3341E 01	-0.5449E 00	-0.2110E-01
0.6930E-01	0.1761E 04	0.3392E 01	-0.4740E-00	-0.
0.2328E-00	-0.1344E-02	0.1761E 04	-0.1609E 01	-0.4740E-00
0.7620E 00	-0.	-0.	0.3889E-01	0.3220E-01
-0.	-0.1478E-01	0.7620E 00	-0.5670E 02	-0.6930E-01

INDEX= 2 (S_2)

COLUMNS 1 THRU 5

-0.1350E-12	-0.3415E 02	0.1760E 04	-0.1583E 01	-0.4740E-00
0.2328E-00	0.5647E 02	0.1761E 04	-0.1609E 01	-0.4740E-00
0.1216E 03	-0.4517E-02	0.5670E 02	-0.5620E-01	-0.1673E-01
-0.	-0.1478E-01	0.7620E 00	-0.2315E-00	-0.6930E-01
-0.1137E-12	0.	-0.1478E-01	0.1216E 03	0.1344E-02

INDEX= 1 (S_1)

COLUMNS 1 THRU 5

-0.1269E-09	0.2673E-14	-0.3415E 02	0.3090E-01	0.9196E-02
0.1216E 03	-0.1216E 03	0.5670E 02	-0.5620E-01	-0.1673E-01
-0.9356E-14	-0.2360E 01	-0.1269E-09	0.5517E-13	0.5389E-13
-0.1137E-12	0.	-0.1478E-01	0.4517E-02	0.1344E-02
0.2791E-12	0.1776E-14	-0.2162E-12	-0.2360E 01	0.7243E-10

INDEX= 0 (S_0) an indication of computation error

COLUMNS 1 THRU 5

0.3029E-09	0.2463E-11	-0.1269E-09	-0.3777E-11	0.4875E-13
-0.9356E-14	0.2988E-09	-0.1269E-09	0.5517E-13	0.5389E-13
-0.8741E-11	0.1853E-15	0.2988E-09	-0.8598E-10	-0.1136E-12
0.2791E-12	0.1776E-14	-0.2162E-12	-0.5333E-09	0.7243E-10
0.1025E-09	-0.3932E-14	0.9130E-12	-0.3166E-06	-0.3671E-09

The very small magnitudes of the elements of the S_0 matrix indicate the benefit of using double precision arithmetic in this subroutine.

OPEN LOOP CHARACTERISTIC EQUATION

COEFFICIENTS OF ASCENDING POWERS OF S (0 TO 5)

0.23598E 01	-0.12164E 03	0.56473E 02	0.17605E 04	0.33922E 01	0.10000E 01

$$\Delta(s) = s^5 + 3.39\,s^4 + 1760.5\,s^3 + 56.47\,s^2 - 121.64\,s + 2.36$$

ROOTS OF OPEN LOOP CHARACTERISTIC EQUATION

REAL	IMAGINARY
0.19691E-01	0.
0.23625E-00	0.
-0.28815E-00	0.
-0.16800E 01	0.41925E 02
-0.16800E 01	-0.41925E 02

7 SIGNIFICANT FIGURES

ELINV (USED IN FILTER PROGRAM)

0.160555E-00	-0.292019E-00	0.940875E-12	0.234736E-01	0.126468E-08
-0.827604E 01	-0.827604E 01	-0.292019E-00	-0.120998E 01	0.234736E-01
-0.368406E-00	-0.	-0.827604E 01	0.562212E 00	-0.120998E 01
-0.	-0.	-0.	0.174600E 02	0.562212E 00
-0.	-0.	-0.	-0.	0.174600E 02

$\text{ELINV} = L^{-1} = [S_1 a, S_2 a, \cdots, S_n a]^*$

This matrix is computed each time but is only used here. With a list of kappa's, the computation and printout done to this point would not be repeated.

PERFORMANCE WEIGHTING FACTORS

DRIFT MINIMIZING (KAPPA-ROOF)	-0.
LOAD MINIMIZING (KAPPA-TILDE)	0.100E 06

WEIGHTED PERFORMANCE INDEX - C

0.1610E 06	-0.	0.	0.5600E 06	0.
-0.	-0.	0.	-0.	0.
0.	0.	0.	0.	0.
0.5600E 06	-0.	0.	0.1610E 06	0.
0.	0.	0.	0.	0.

$$C = \hat{\kappa}\hat{C} + \tilde{\kappa}\tilde{C}$$

SUBROUTINE POLYWG PRINTS COEFFICIENTS OF EVEN POWERS OF S (ODD POWERS ARE ZERO), N ROOTS OF POLYNOMIAL FORMED BY EVEN COEFS. ONLY, N L.H.P. ROOTS OF 2NTH ORDER POLYNOMIAL, AND THE COEFFICIENTS GENERATED BY THE L.H.P. ROOTS.

POWER OF S	COEFFICIENT
0	0.84655929E 04
2	-0.22416455E 08
4	0.16891450E 09
6	-0.52179975E 08
8	-0.35095707E 04
10	-0.09999999E 01

ROOTS FROM EVEN COEFFICIENTS

0.3787315E-03	0.
0.1382508E-00	0.
0.3097846E 01	0.
-0.1756404E 04	0.7007595E 04
-0.1756404E 04	-0.7007595E 04

ROOTS WITH NEGATIVE REAL PARTS

-0.1946103E-01	-0.
-0.3718209E-00	-0.
-0.1760070E 01	-0.
-0.5228745E 02	-0.6701030E 02
-0.5228745E 02	0.6701030E 02

COEFFICIENTS OF POLYNOMIAL GENERATED BY ROOTS IN LEFT HALF PLANE

POWER OF S	COEFFICIENT REAL	COEFFICIENT IMAGINARY
5	0.09999999E 01	0.
4	0.10672624E 03	-0.
3	0.74500308E 04	0.15258789E-04
2	0.15614923E 05	-0.10967255E-04
1	0.50289033E 04	-0.18477440E-04
0	0.92008653E 02	-0.

CONTROLLABILITY MATRIX (Actually the transpose, D*)

0.	0.	0.	0.	0.174600E 02
-0.	-0.	-0.	0.174600E 02	-0.586656E 02
-0.368406E-00	0.	-0.827604E 01	-0.586656E 02	-0.305412E 05
-0.702633E 01	-0.827604E 01	0.277820E 02	-0.305412E 05	0.205899E 06
0.672588E 03	0.277820E 02	0.144760E 05	0.205899E 06	0.530760E 08

The determinant of this matrix must be non-zero if the system is to be controlled. If it were zero, an attempt would still be made to compute the control vector, but its terms would be the computational equivalent of infinity.

S-MATRICES OF CLOSED LOOP

INDEX= 4

COLUMNS 1 THRU 5

0.1067E 03	-0.1940E-01	0.1000E 01	-0.2110E-01	-0.
-0.	0.1067E 03	0.1000E 01	-0.	-0.
0.6930E-01	-0.	0.1067E 03	-0.4740E-00	-0.
0.	0.	0.	0.1067E 03	0.1000E 01
0.3649E 04	0.6909E 04	0.3229E 05	-0.7446E 04	0.3220E-01

INDEX= 3

COLUMNS 1 THRU 5

0.7446E 04	-0.2070E 01	0.1067E 03	-0.2725E 01	-0.2110E-01
0.6930E-01	0.7450E 04	0.1067E 03	-0.4740E-00	-0.
0.7394E 01	-0.1344E-02	0.7450E 04	-0.5059E 02	-0.4740E-00
0.3649E 04	0.6909E 04	0.3229E 05	0.3366E 01	0.3220E-01
0.2238E 04	0.1517E 03	0.1160E 05	-0.1562E 05	-0.6930E-01

INDEX= 2

COLUMNS 1 THRU 5

0.1531E 05	-0.2902E 03	0.6763E 04	-0.5056E 02	-0.4740E-00
0.7394E 01	0.1561E 05	0.7450E 04	-0.5059E 02	-0.4740E-00
-0.1214E 04	-0.3275E 04	0.3168E 03	-0.1784E 01	-0.1673E-01
0.2238E 04	0.1517E 03	0.1160E 05	-0.7393E 01	-0.6930E-01
0.4788E 03	-0.5222E 03	0.1517E 03	-0.5029E 04	0.1344E-02

INDEX= 1

COLUMNS 1 THRU 5

0.3275E 04	-0.3572E 04	-0.2902E 03	0.9811E 00	0.9196E-02
-0.1214E 04	0.1754E 04	0.3168E 03	-0.1784E 01	-0.1673E-01
0.3183E-11	-0.9201E 02	-0.4184E-09	-0.1345E-10	0.2155E-12
0.4788E 03	-0.5222E 03	0.1517E 03	0.1434E-00	0.1344E-02
0.1179E-06	-0.1199E-06	0.2679E-05	-0.9201E 02	0.2014E-08

INDEX= 0

COLUMNS 1 THRU 5

0.7800E-07	0.3103E-10	0.1353E-09	0.9819E-11	0.4219E-13
0.3183E-11	0.7795E-07	-0.4184E-09	-0.1345E-10	0.2155E-12
0.7369E-09	0.4017E-09	0.9034E-07	0.5229E-09	-0.3887E-11
0.1179E-06	-0.1199E-06	0.2679E-05	-0.5528E-07	0.2014E-08
-0.2169E-05	0.1579E-04	-0.1051E-03	0.2200E-04	-0.1910E-06

A-TILDE-TRANSPOSE (ATT USED IN FILTER PROGRAM) $(\tilde{A}^*)$

-0.322000E-01	-0.	0.693000E-01	-0.	0.364936E 04
-0.194000E-01	-0.	-0.	-0.	0.690904E 04
0.100000E 01	0.100000E 01	-0.	-0.	0.322896E 05
-0.211000E-01	-0.	-0.474000E-00	-0.	-0.744631E 04
-0.	-0.	-0.	0.100000E 01	-0.106694E 03

OPTIMAL CLOSED-LOOP ROOTS

REAL	IMAGINARY
-0.19461E-01	-0.
-0.37182E-00	-0.
-0.17601E 01	-0.
-0.52287E 02	-0.67010E 02
-0.52287E 02	0.67010E 02

(Same as "roots with negative real parts." Restated for convenience.)

COMPUTED FEEDBACK CONTROL VECTOR (TERMS 1 THRU N) (g)

0.20897E 03
0.39571E 03
0.18493E 04
-0.32565E 03
-0.59184E 01

$$\psi = 208.97\,\alpha + 395.71\,\phi + 1849.3\,\dot{\phi} - 325.65\,\beta - 5.92\,\dot{\beta}$$

CLOSED LOOP ROOTS ACHIEVED BY CONTROL VECTOR

-0.19461E-01	0.
-0.37182E-00	0.
-0.17601E 01	0.
-0.52288E 02	0.67007E 02
-0.52288E 02	-0.67007E 02

These roots should match the optimal closed loop roots above. Roundoff error accounts for the differences.

DESIGN FILTER TO APPROXIMATE DESIRED SYSTEM POLE CONFIGURATION

A-TILDE-TRANSPOSE MATRIX ($\tilde{A}^*$, from CNTRL2)

-0.032200	-0.	0.069300	-0.	3649.359985
-0.019400	-0.	-0.	-0.	5909.049988
1.000000	1.000000	-0.	-0.	32289.699951
-0.021100	-0.	-0.474000	-0.	-7446.500000
0.	-0.	-0.	1.000000	-106.688999

G-LITTLE (CONTROL VECTOR) (from CNTRL2

208.969999
395.709999
1849.399994
-325.660000
-5.918000

HSTAR MATRIX

SENSOR 1	-0.	1.000	-0.	-0.	-0.	(Senses ϕ)
SENSOR 2	-0.	-0.	-0.	1.000	-0.	(Senses β)

COEFFICIENTS OF OPEN LOOP CHARACTERISTIC EQUATION (POWERS OF S FROM 0 TO N)

0.23598E 01	-0.12164E 03	0.56473E 02	0.17605E 04	0.33922E 01	0.10000E 01

ELINV MATRIX (L^{-1} from CNTRL2)

0.1605550E-00	-0.2920190E-00	0.9409000E-12	0.2347360E-01	0.1264700E-08
-0.8276040E 01	-0.8276040E 01	-0.2920190E-00	-0.1209980E 01	0.2347360E-01
-0.3684060E-00	-0.	-0.8276040E 01	0.5622120E 00	-0.1209980E 01
0.	-0.	-0.	0.1746000E 02	0.5622120E 00
0.	-0.	-0.	-0.	0.1746000E 02

OBSERVABILITY MATRIX

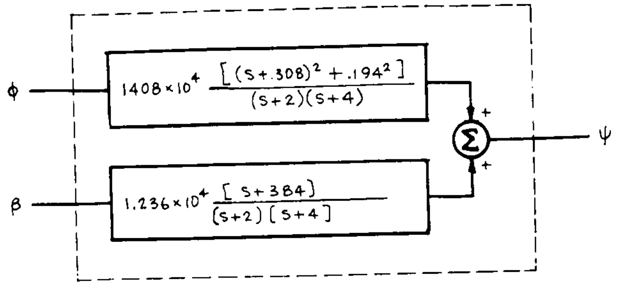
-0.	-0.	-0.	-0.	0.693000E-01
0.100000E 01	-0.	-0.	-0.	0.
-0.	-0.	0.100000E 01	-0.	0.
-0.	0.100000E 01	0.	0.	-0.474000E-00
-0.	-0.	0.	0.100000E 01	-0.

If formation of this matrix is possible, (the subroutine OBSMAT checks column independence during formation) the system is observable and a realizable filter can be derived.

FILTER PARAMETERS

2 - PORT FILTER

POLYNOMIAL COEFFICIENTS IN ASCENDING POWERS OF S (0 TO 2)



NUMERATOR NO. 1

0.1871E 07 0.8681E 07 0.1408E 08

NUMERATOR NO. 2

0.4746E 07 0.1236E 05 0.

COMMON DENOMINATOR

0.8000E 01 0.6000E 01 0.1000E 01

FILTER ROOTS

NUMERATOR NO. 1

	-0.308200E-00	0.194485E-00
	-0.308200E-00	-0.194485E-00

The single root in Numerator No. 1 is obtainable directly from the coefficients above, hence it is not computed or printed.

COMMON DENOMINATOR

	-0.200000E 01	-0.
	-0.400000E 01	-0.

RESULTING POLE POSITIONS

-0.194610E-01	0.
-0.371820E-00	0.
-0.176009E 01	0.
0.466643E 02	0.485572E 02
0.466643E 02	-0.485572E 02
-0.522848E 02	0.670110E 02
-0.522848E 02	-0.670110E 02

All poles except those added by the filter should be in the same positions as if all variables had been measured.

These poles should be near the filter poles. The fact that they are not, and that they are unstable, indicates that the filter poles should be chosen further from the jω axis.

FILTER PARAMETERS

2 - PORT FILTER

POLYNOMIAL COEFFICIENTS IN ASCENDING POWERS OF S (0 TO 2)

NUMERATOR NO. 1

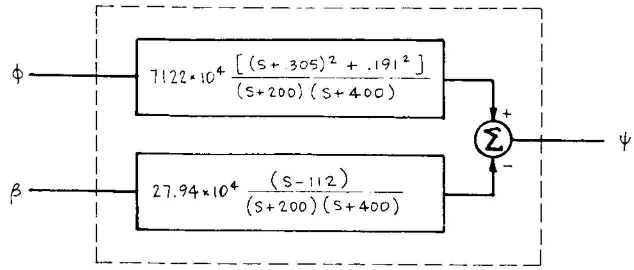
0.9234E 07   0.4347E 08   0.7122E 08

NUMERATOR NO. 2

0.3143E 08 -0.2794E 06   0.

COMMON DENOMINATOR

0.8000E 05   0.6000E 03   0.1000E 01



FILTER ROOTS

NUMERATOR NO. 1

-0.305183E-00	0.191114E-00
-0.305183E-00	-0.191114E-00

COMMON DENOMINATOR

-0.200000E 03	-0.
-0.400000E 03	-0.

RESULTING POLE POSITIONS

-0.194610E-01	0.
-0.371781E-00	0.
-0.176034E 01	0.
-0.525566E 02	0.
-0.444116E 03	0.
-0.522842E 02	0.670105E 02
-0.522842E 02	-0.670105E 02

These two poles are stable and nearer to the filter poles, demonstrating the asymptoticity property of the design procedure.

6. RESULTS OF COMPUTER FLIGHTS

The control system synthesis and the simulated flights were accomplished for a five-dimensional model based on the rigid-body motions with perfect sensors and actuator. The equations of motion for this model are*

$$\dot{\alpha} = -0.0322\alpha - 0.0194\phi + \dot{\phi} - 0.0211\beta \tag{6-1}$$

$$\ddot{\phi} = 0.0693\alpha - 0.474\beta \tag{6-2}$$

$$\ddot{\beta} = 0.762\alpha - 1760.5\beta - 3.36\dot{\beta} + 17.5\psi \tag{6-3}$$

The two quantities, besides control, appearing in the criterion integrand are drift acceleration

$$\ddot{z} = \frac{T + T_c - D}{m}\phi + \frac{N_\alpha}{m}\alpha + \frac{T_c}{m}\beta = q_1 \cdot x^{**} \tag{6-4}$$

and bending load

$$L = 1.26\alpha + 4.4\beta = q_2 \cdot x^{**} \tag{6-5}$$

For the flights, lateral drift was actually computed by integrating

$$\dot{z} = V(\phi + \alpha_w - \alpha) \ . \tag{6-6}$$

It will help the interpretation of the results of the simulated flights to review the design procedure very quickly. The control is of the form

$$\psi = g \cdot x = g_1\alpha + g_2\phi + g_3\dot{\phi} + g_4\beta + g_5\dot{\beta} \ , \tag{6-7}$$

---

*Numerical values were derived from NASA-supplied documents.
**See Section 3.

where the g_i are determined using the design programs so as to minimize the criterion integral

$$\int_0^{\infty}\left(\hat{\kappa} z^2 + \tilde{\kappa}\left(\frac{L}{10^7}\right)^2 + \psi^2\right)dt \quad , \tag{6-8}$$

where $\hat{\kappa}$ and $\tilde{\kappa}$ are weighting factors chosen by the designer. The following restrictions should be kept in mind: (6-8) is minimum for initial condition errors in the absence of winds; because of the term ψ^2 (required for stability) there is always some amount of drift minimization and load minimization called for, since ψ is a linear combination of all the states while drift and load are linear combinations of states, also. Because of the latter restriction, $\hat{\kappa}$ and $\tilde{\kappa}$ are not directly related to the minimization of drift and load, respectively, in a straightforward manner. Indeed, the control can be written in the form

$$\psi = \hat{a}\ddot{z} + \tilde{a} L + \psi' \quad , \tag{6-9}$$

so that (6-8) becomes

$$\int_0^{\infty}\left[(\hat{\kappa} + \hat{a})\ddot{z}^2 + 2\hat{a}\tilde{a}\ddot{z}L + \left(\frac{\tilde{\kappa}}{10^{14}} + a\right)L^2 + 2(\hat{a}\ddot{z} + \tilde{a}L)\psi' + \psi'^2\right]dt \quad , \tag{6-10}$$

where $\hat{a}$, $\tilde{a}$, and ψ' depend implicitly upon $\hat{\kappa}$ and $\tilde{\kappa}$.

The optimal control design was carried out for a wide range of values of $\hat{\kappa}$ and $\tilde{\kappa}$. The resulting control systems were "flown" in the five-dimensional model of (6-1), (6-2), and (6-3) for various conditions, i.e., no wind and the wind shown in Figure 6-1, linear control ($\psi = g \cdot x$) and bang-bang control ($\psi = \text{sgn}\, g \cdot x$), small initial offset ($\alpha = 0.1°$, $\phi = 0.5°$) and large initial offsets ($\alpha = 1°$, $\phi = 5°$). Figures 6-2 through 6-9 are typical of the computer output.* The simulations were run on the IBM 7094 of the Hughes Scientific Computing Department. The

*In Figures 2-9 K1 = $\hat{\kappa}$ and K2 = $\tilde{\kappa}$.

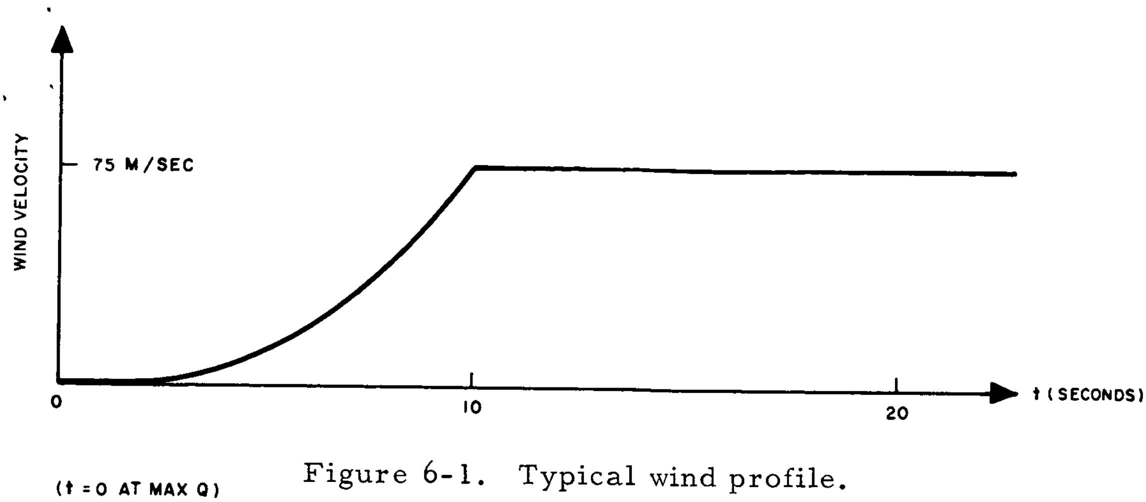


Figure 6-1. Typical wind profile.

program was coded in Fortran IV and employed a straightforward differential equation integration scheme available as a library subroutine.

One fact was evident from the computer traces — all the controllers exhibited the characteristics of "minimum drift control" in the sense of Reference 1; i.e., "$\dot{z}$ goes to zero as soon as the transient oscillation around the center of gravity dies out." *In fact, $\dot{z}$ goes to zero as soon as the wind velocity becomes constant. "Minimum drift" feedback gains in the sense of Reference 1 are never obtained, since all the states are fed back. Another fact is also easily deduced from the computer results — the drift performance is rather insensitive to $\hat{\kappa}$ and $\tilde{\kappa}$. Figure 6-10 shows drift as a function of these quantities for linear control and small initial offsets (for the larger initial offsets just multiply by 5). Note that drift decreases slightly as $\hat{\kappa}$ increases, as expected. For the runs with wind present, the differences in drift performance are not readable from the output plots; the common value for linear control and small offsets is 800 meters. It is also interesting that both linear and bang-bang control resulted in nearly identical drift performance for all the conditions investigated.

The peak bending load is much more sensitive to parameter and control changes. Figure 6-11 shows peak load for the cases used for Figure 6-6. Note that L has a maximum between $\tilde{\kappa} = 0$ and $\tilde{\kappa} = 10^5$. It is obvious from (6-10) that the criterion is more sensitive to $\tilde{a}$ than $\tilde{\kappa}$ for small values of $\tilde{\kappa}$; it is not surprising, then, that the peak load is not

*Reference 1. E.D. Geissler, "Problems in Attitude Stabilization of Large Guided Missiles," Aerospace Engineering, Oct. 1960, p. 24.

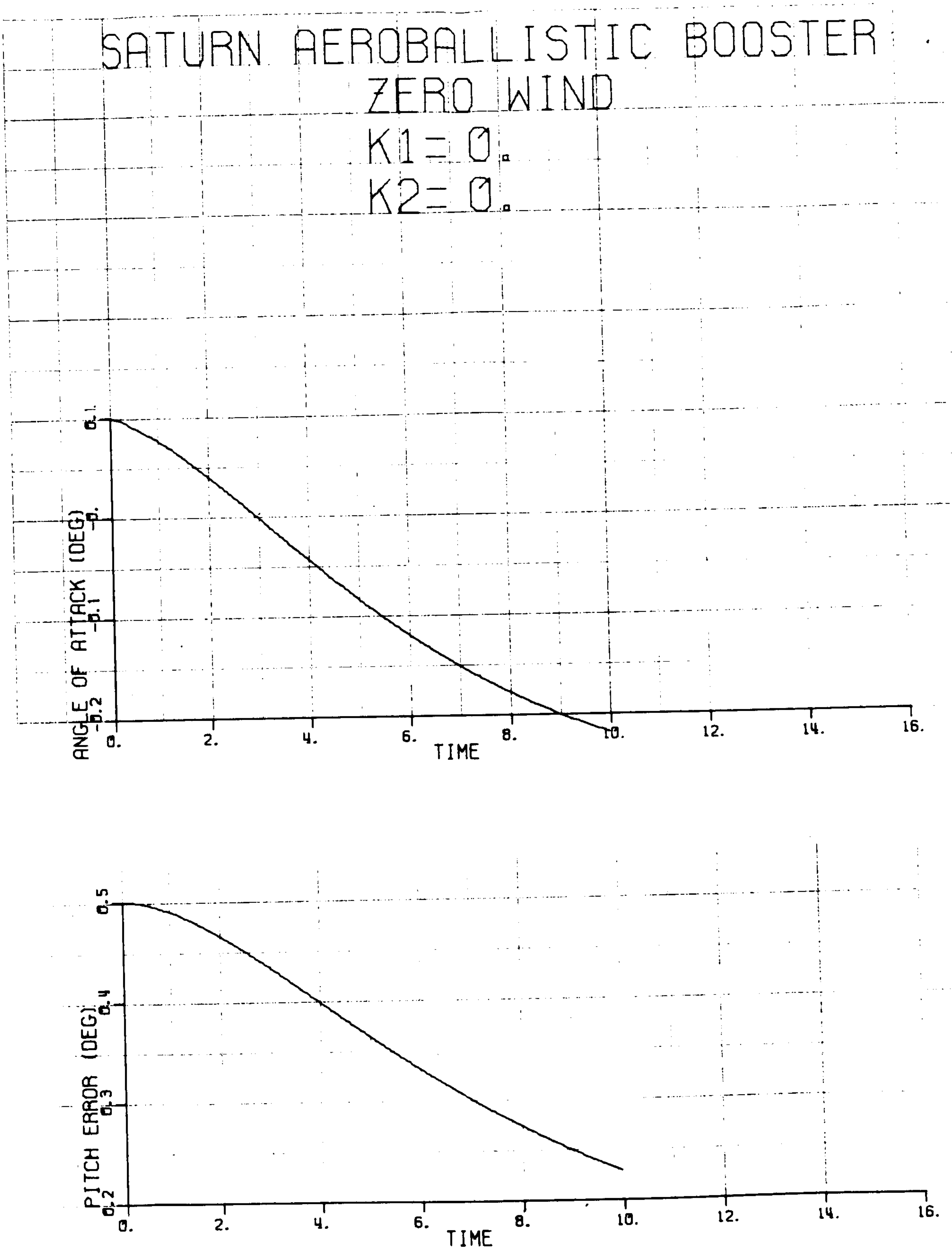


Figure 6-2. Linear control.

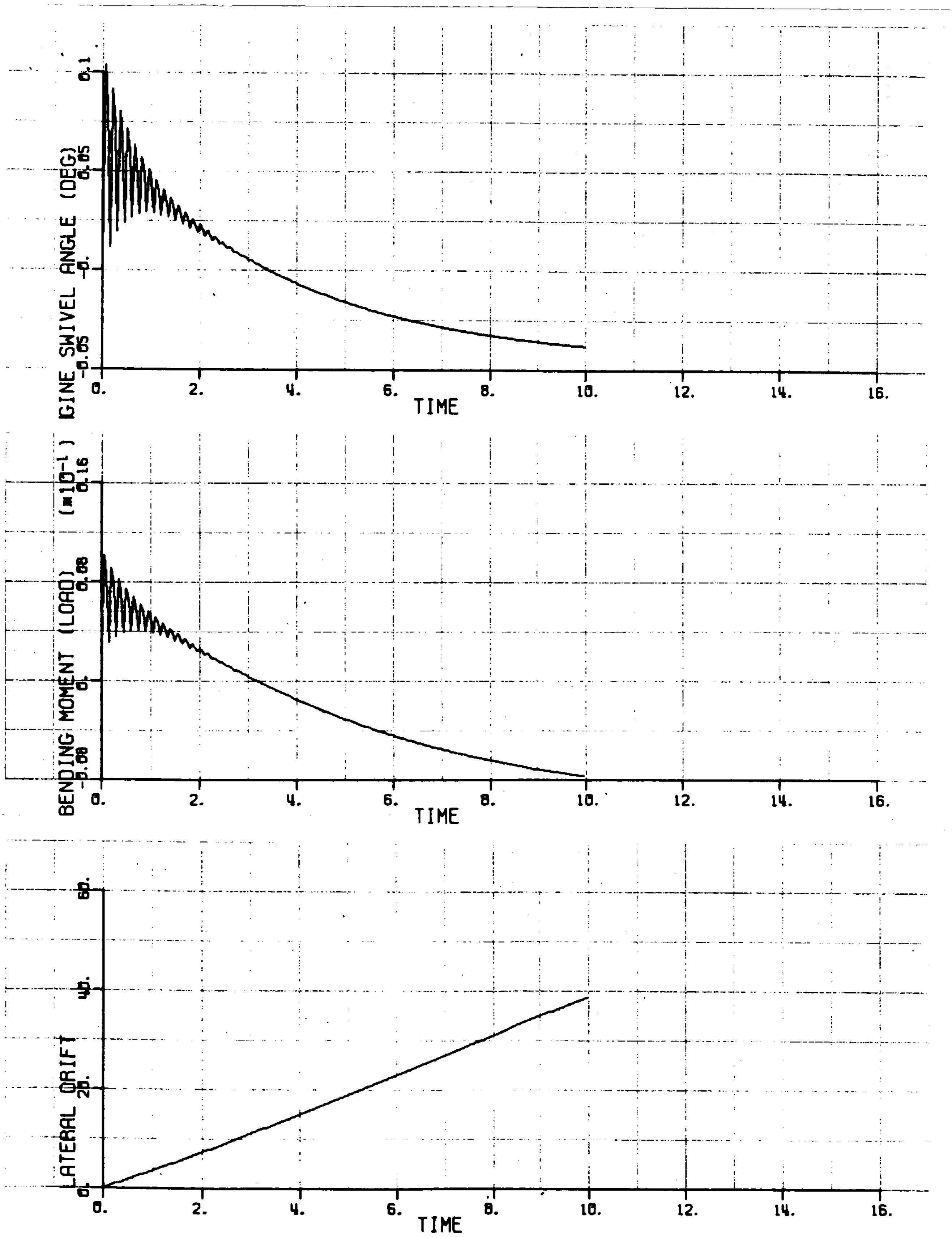


Figure 6-2. Linear control (continued).

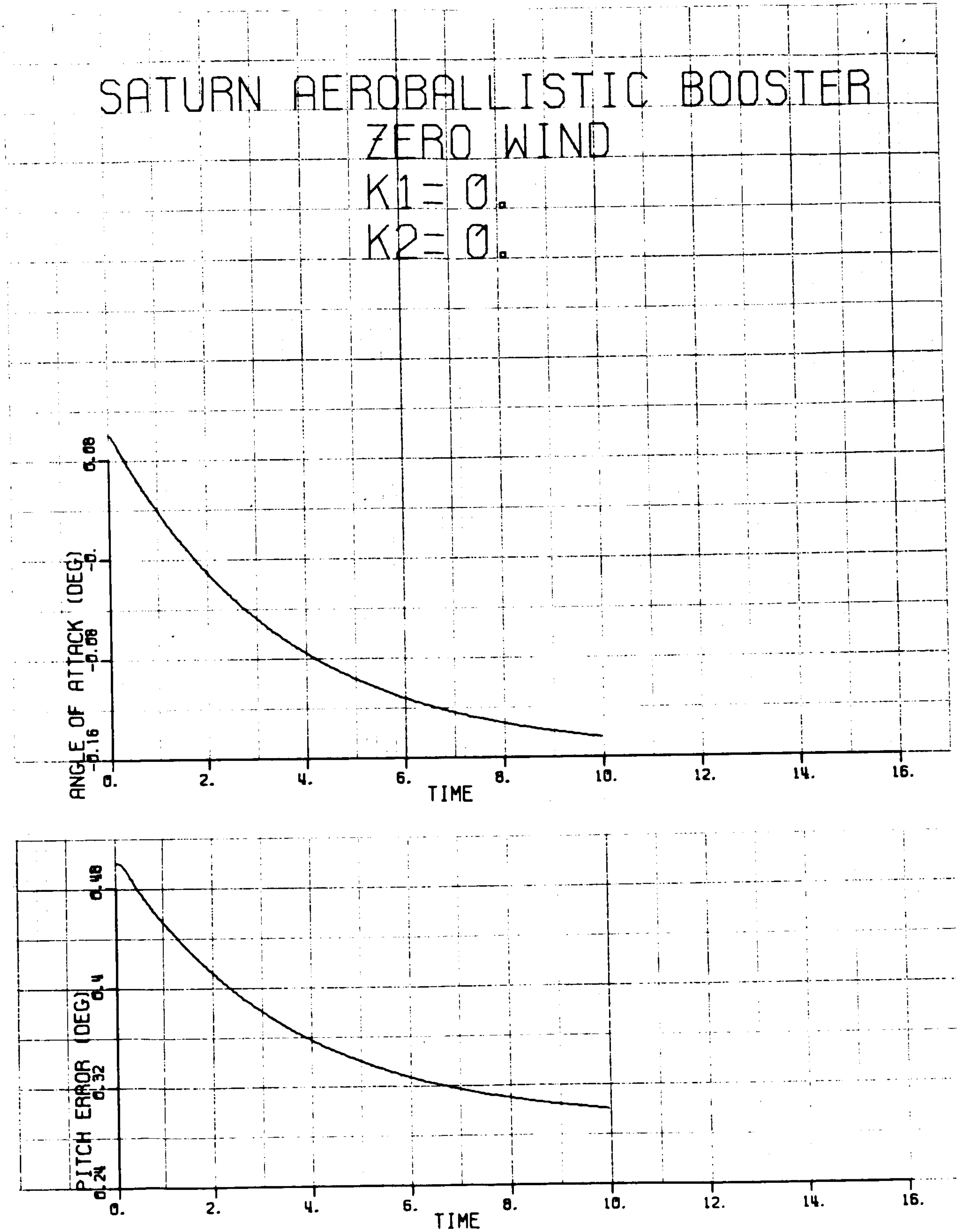


Figure 6-3. Bang-bang control.

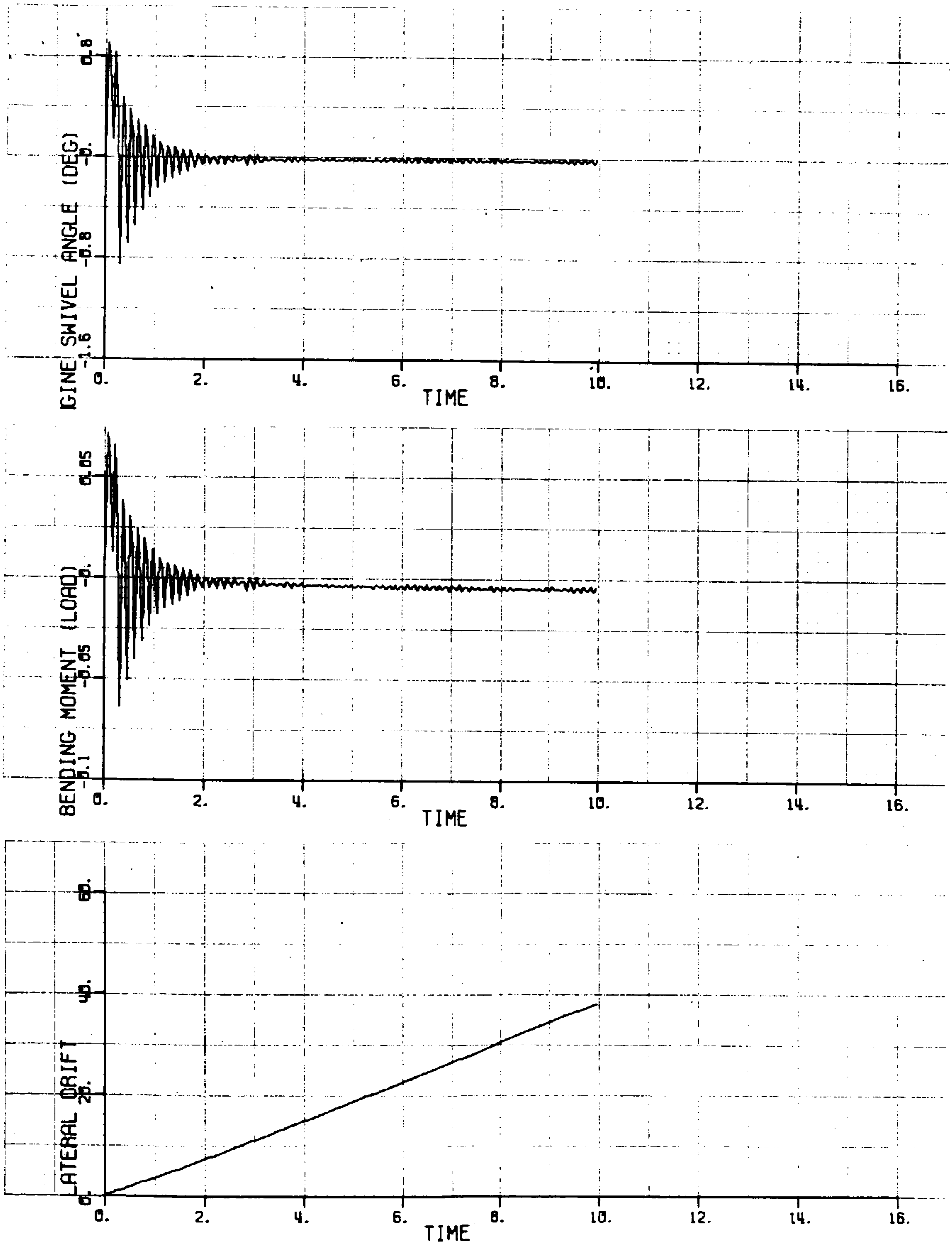


Figure 6-3. Bang-bang control (continued).

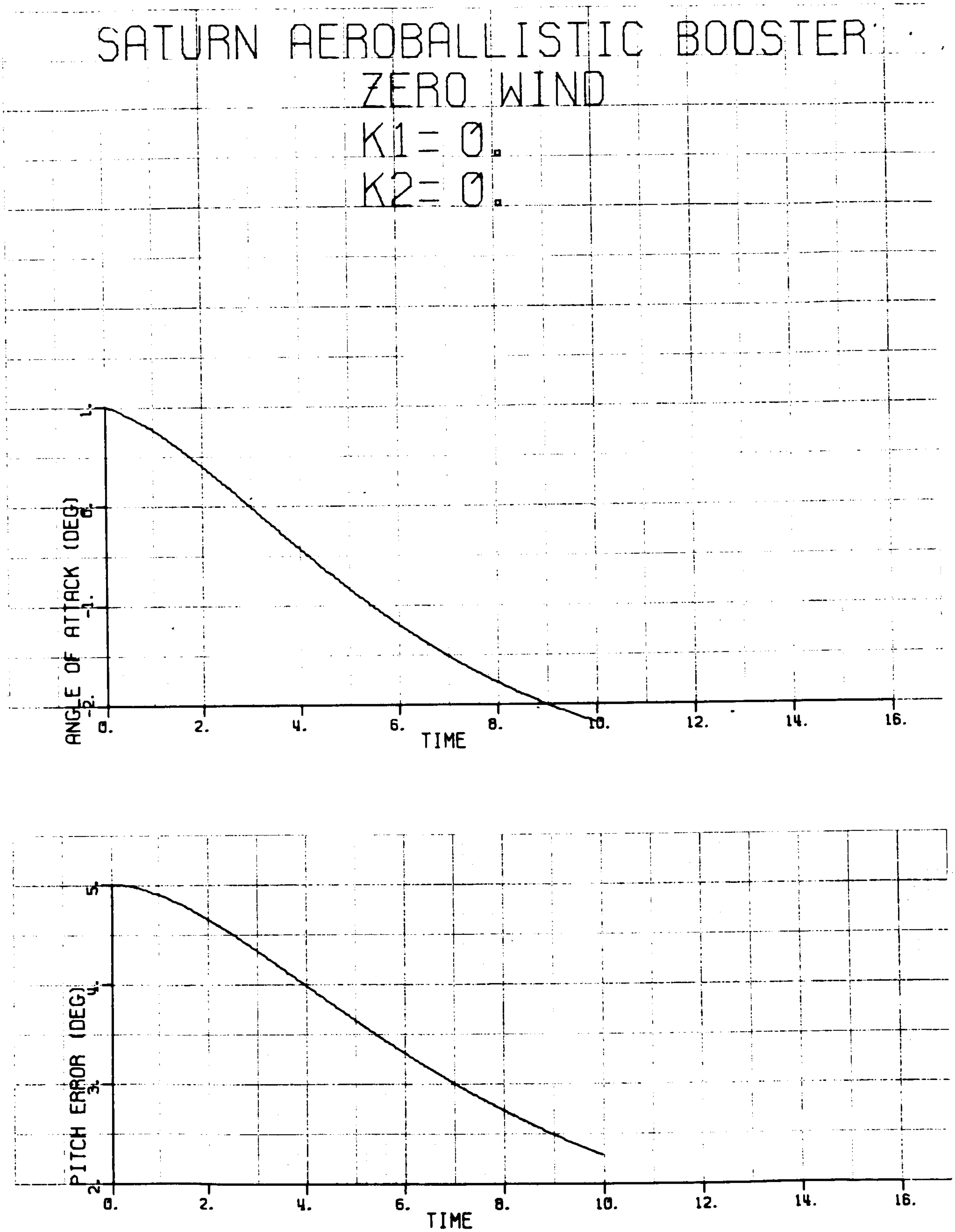


Figure 6-4. Linear control.

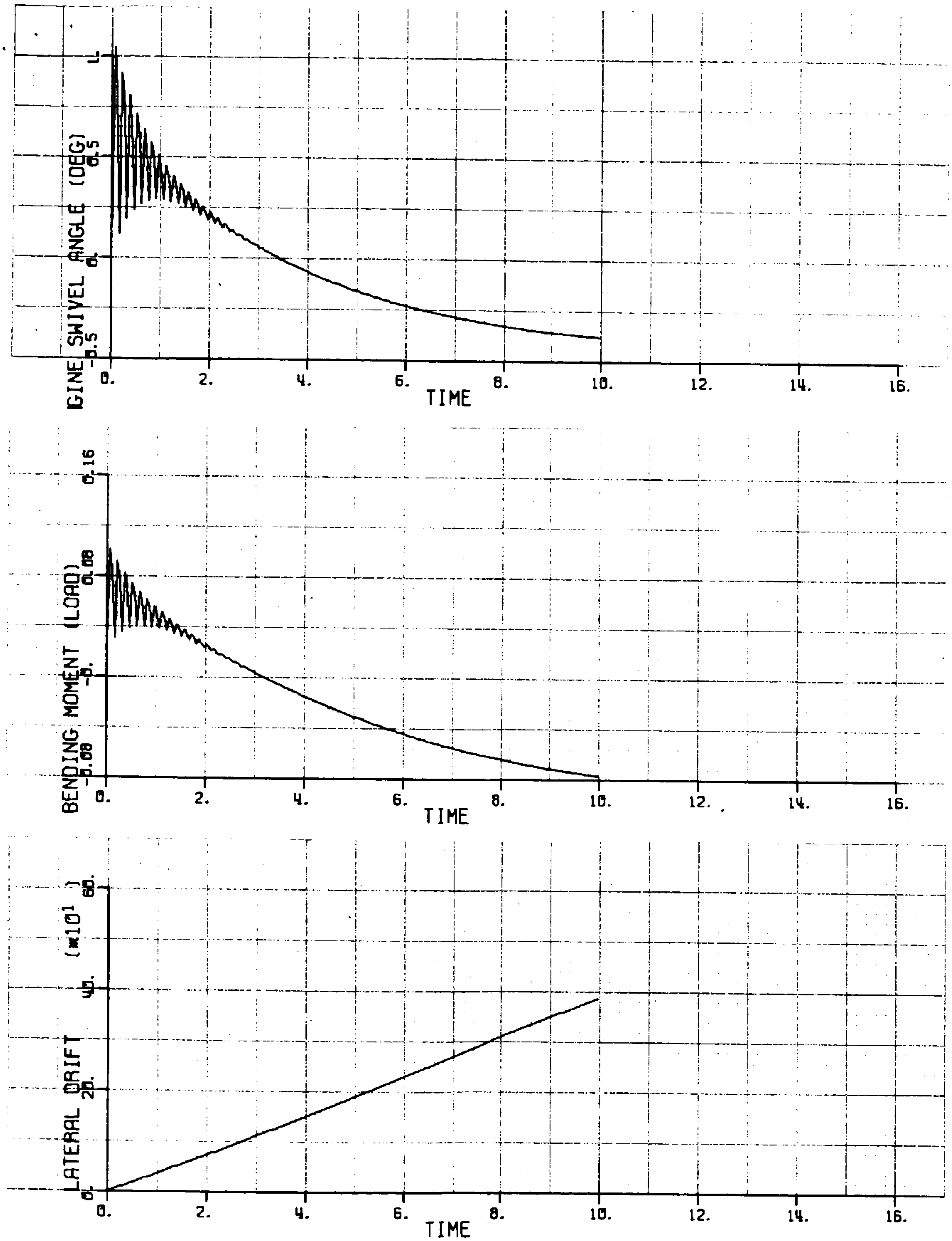


Figure 6-4. Linear control (continued).

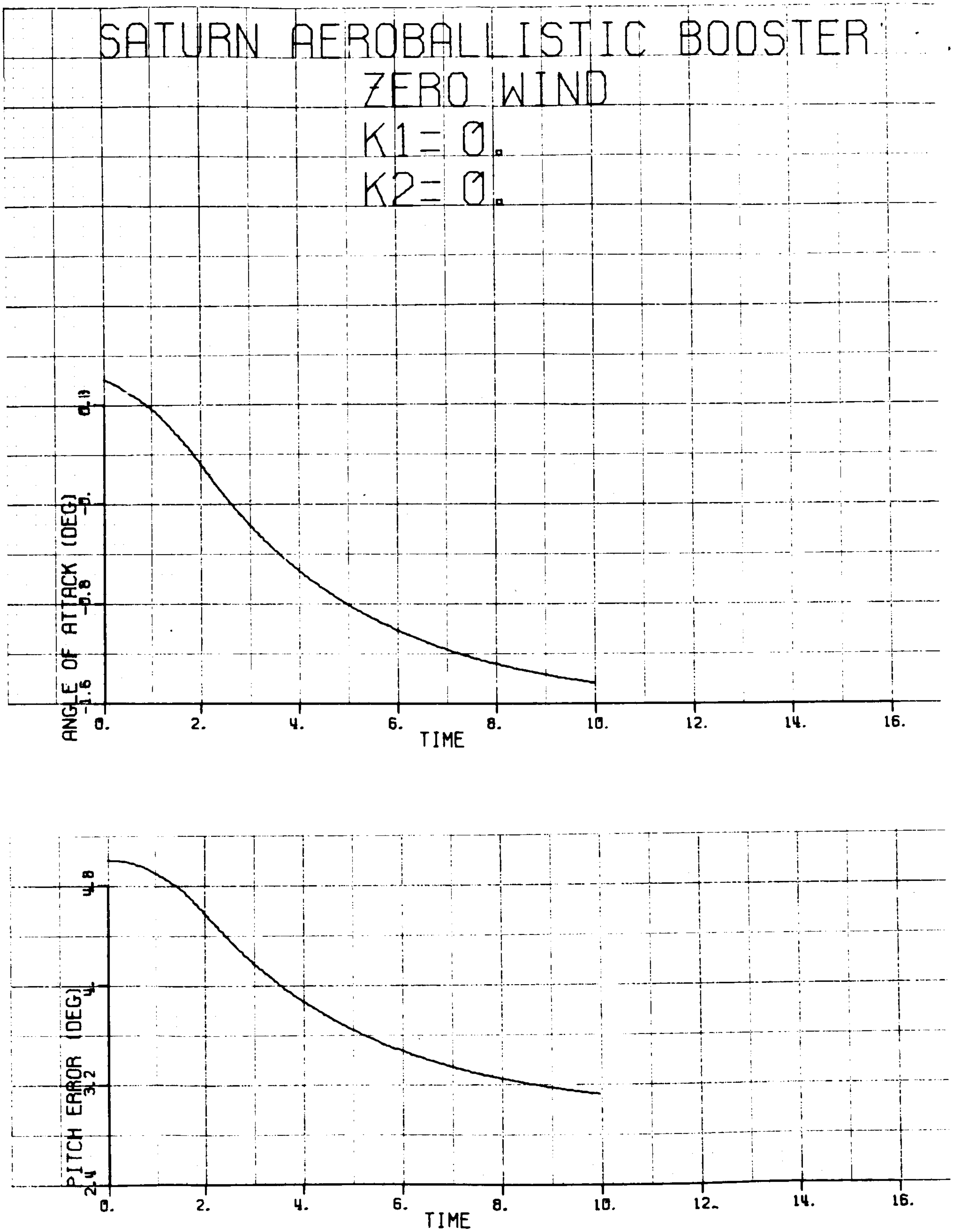


Figure 6-5. Bang-bang control.

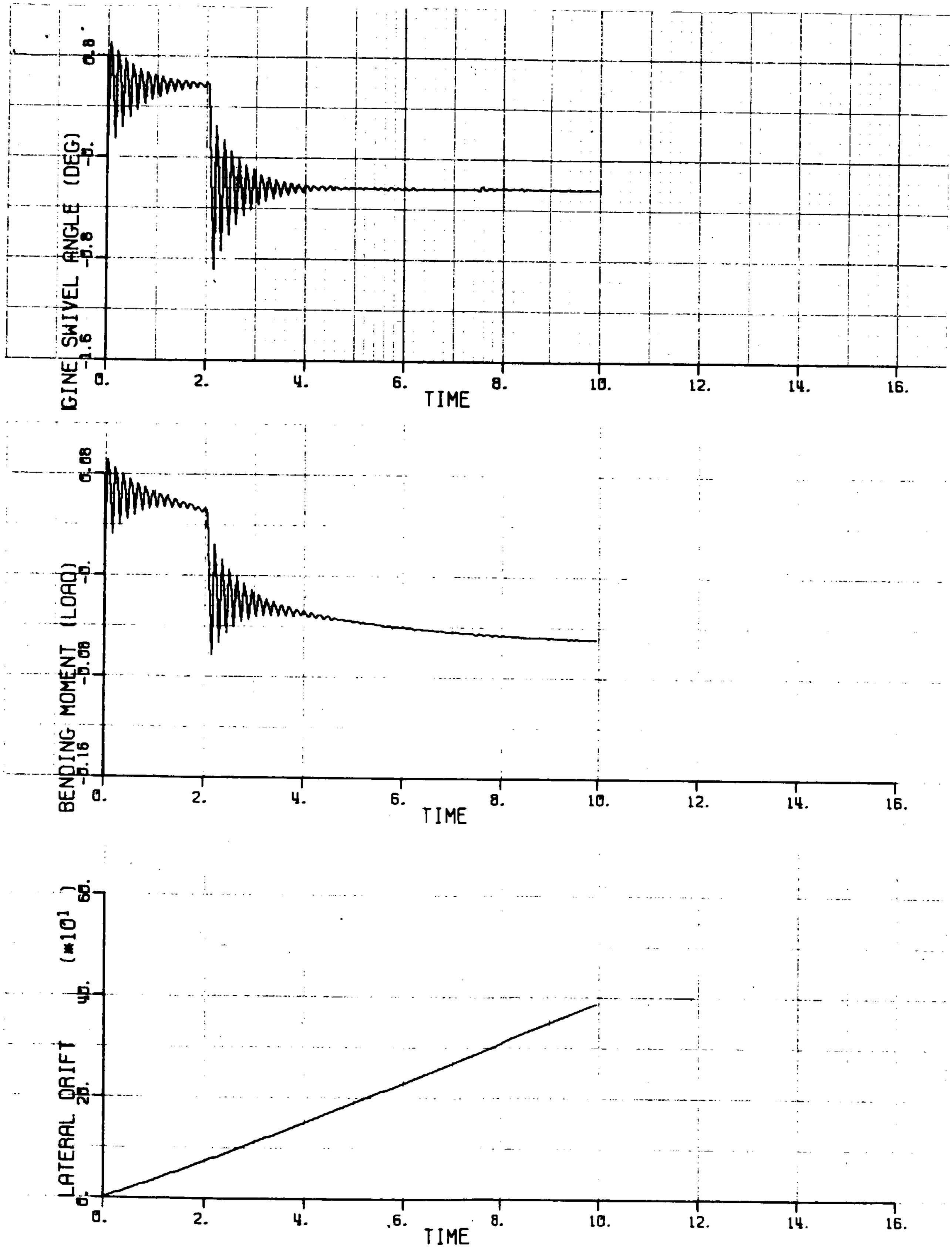


Figure 6-5. Bang-bang control (continued).

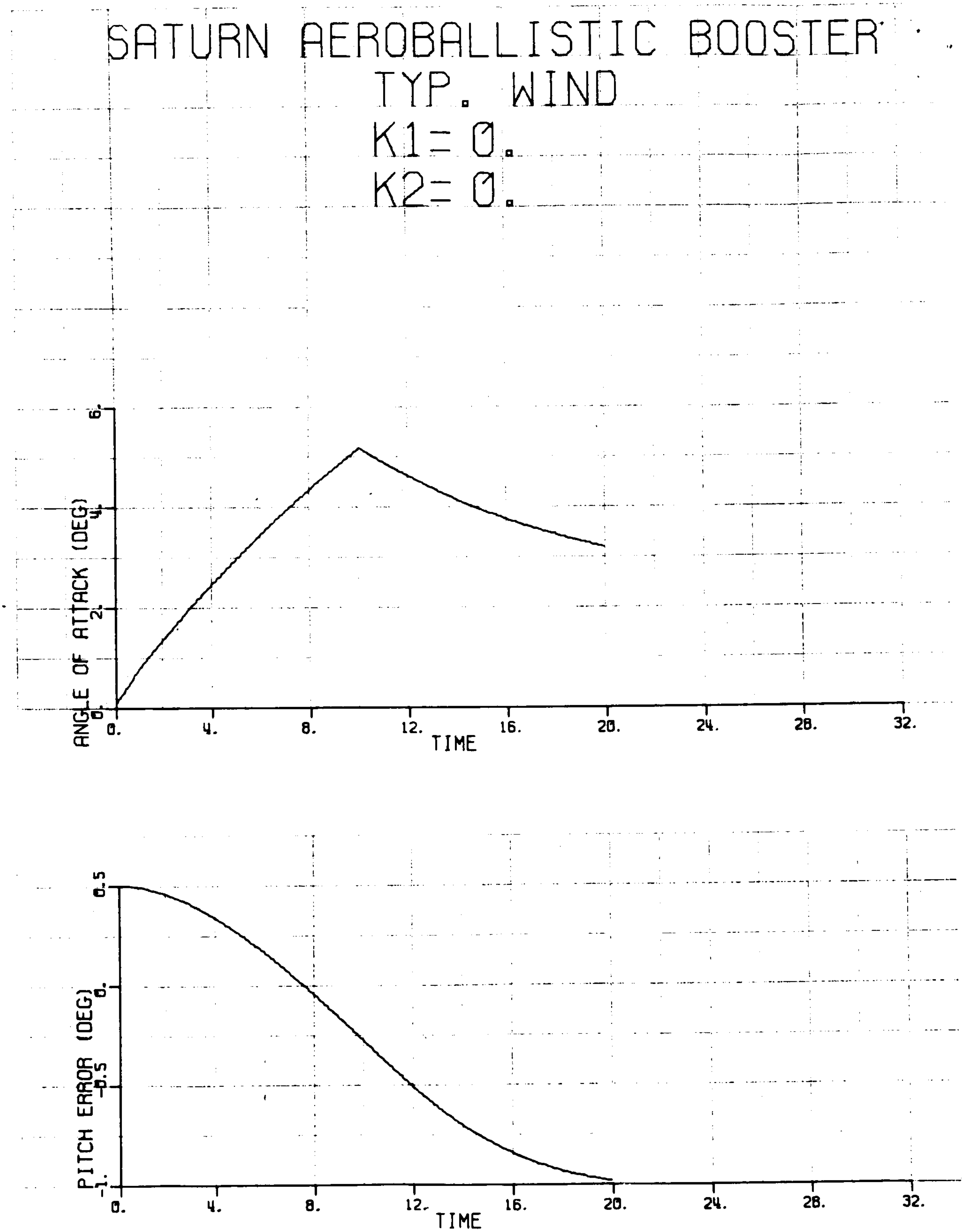


Figure 6-6. Linear control.

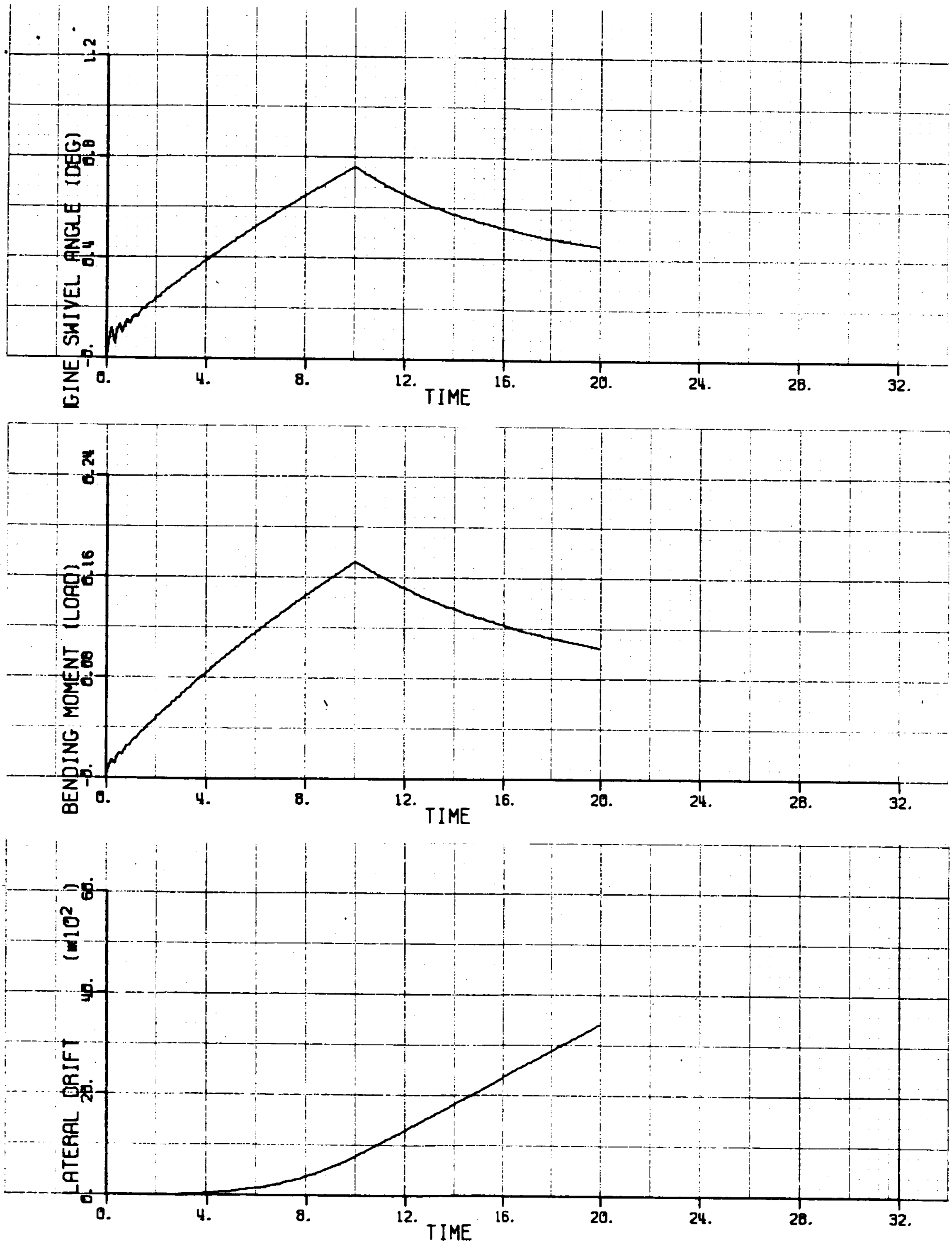


Figure 6-6. Linear control (continued).

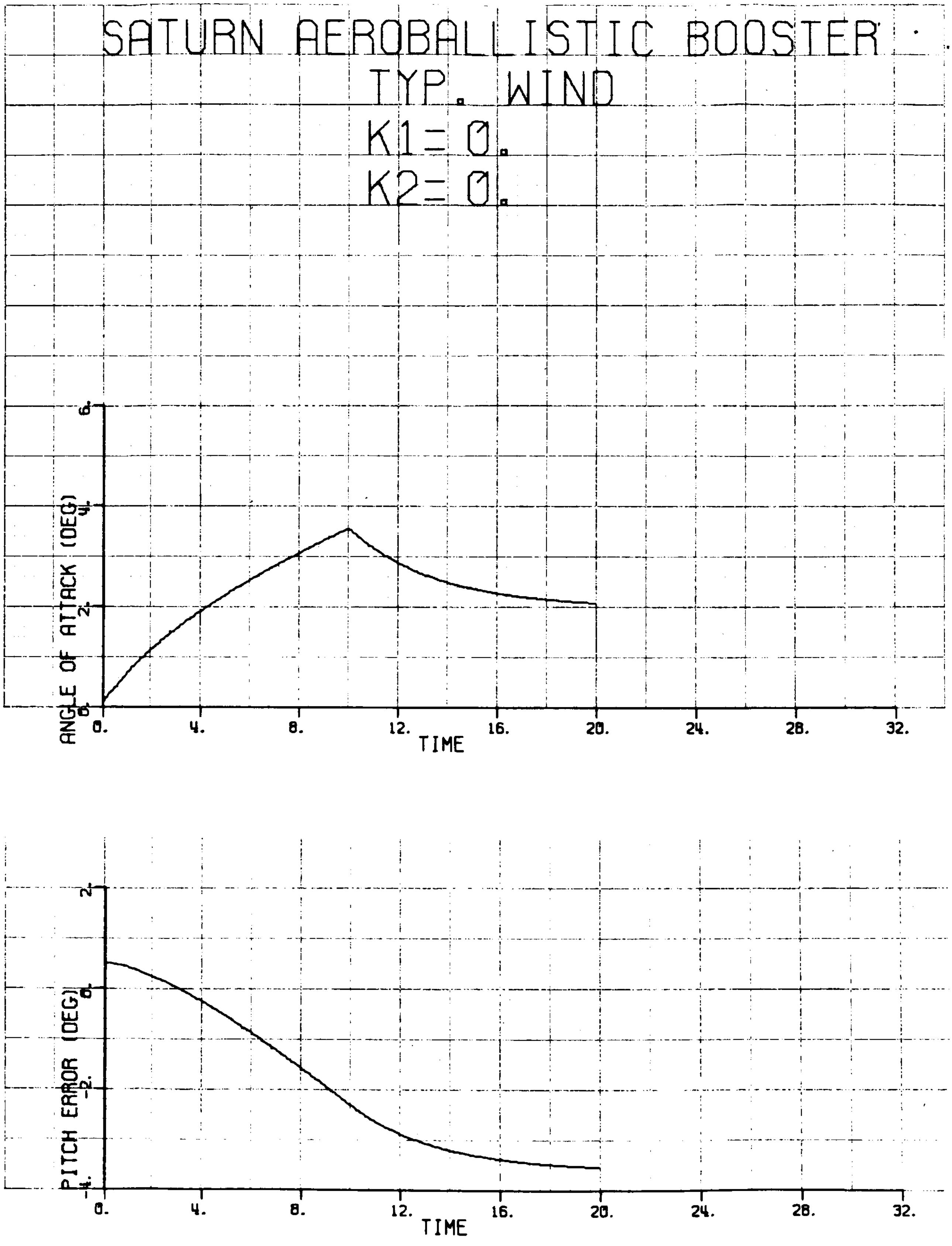


Figure 6-7. Bang-bang control.

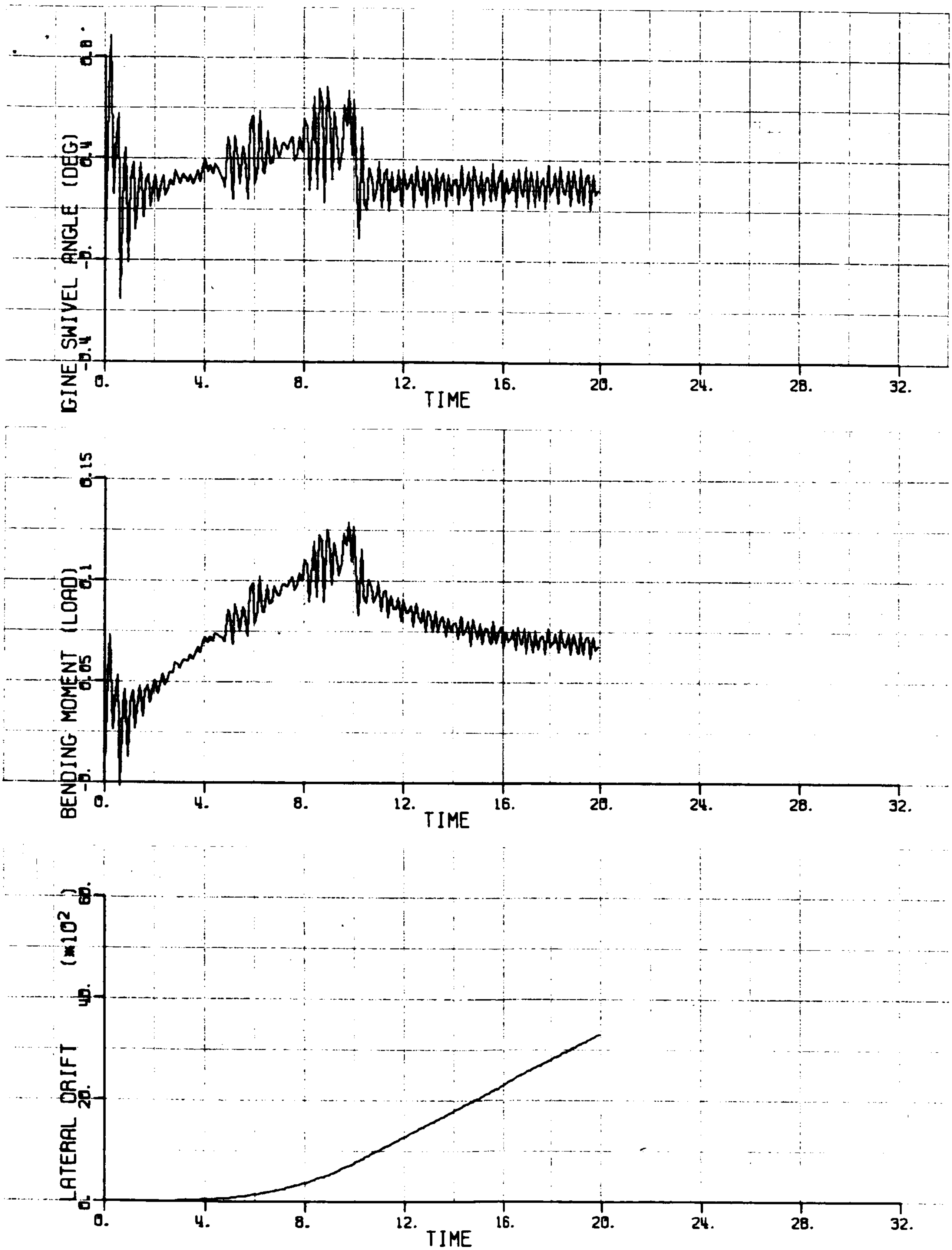


Figure 6-7. Bang-bang control (continued).

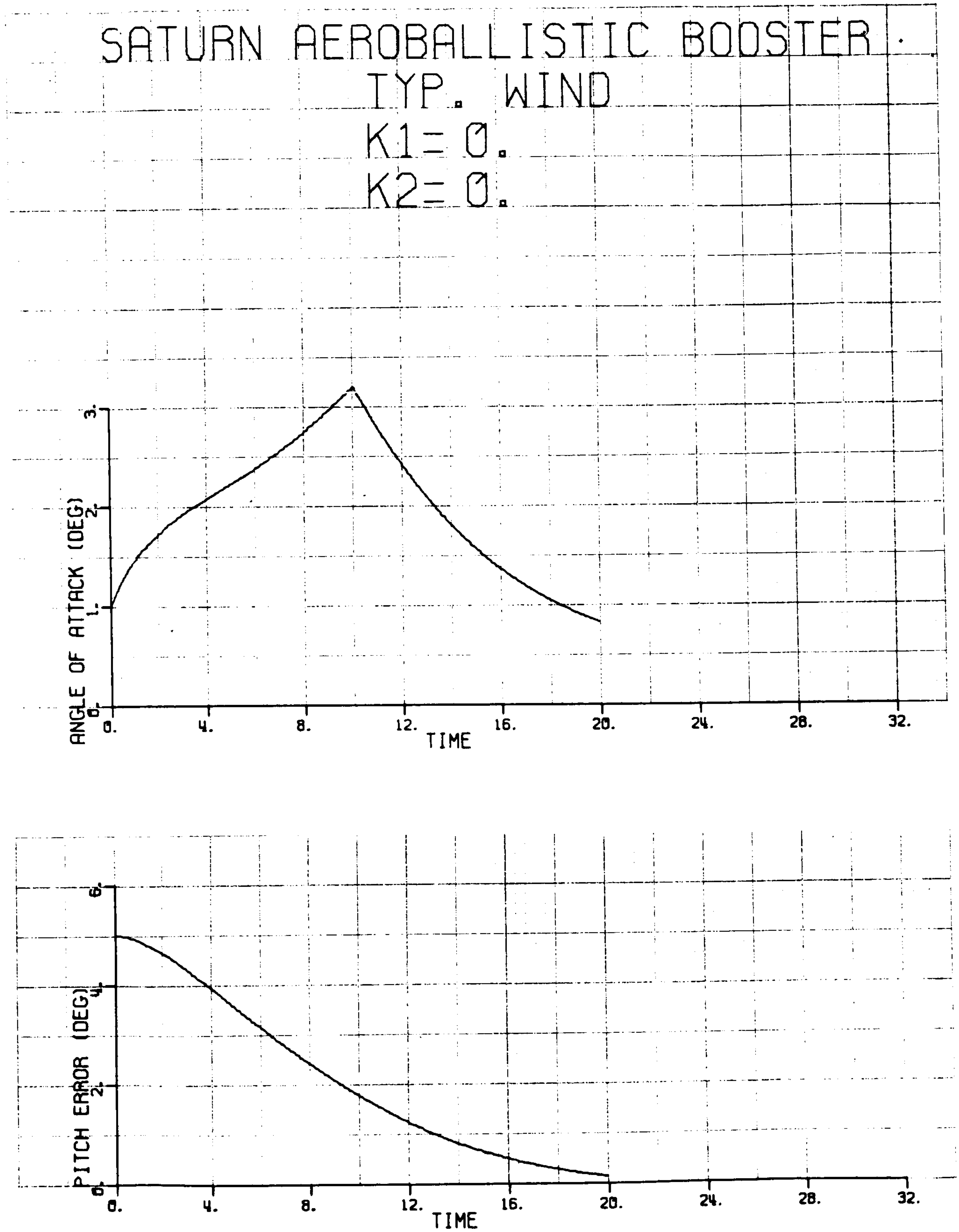


Figure 6-8. Linear control.

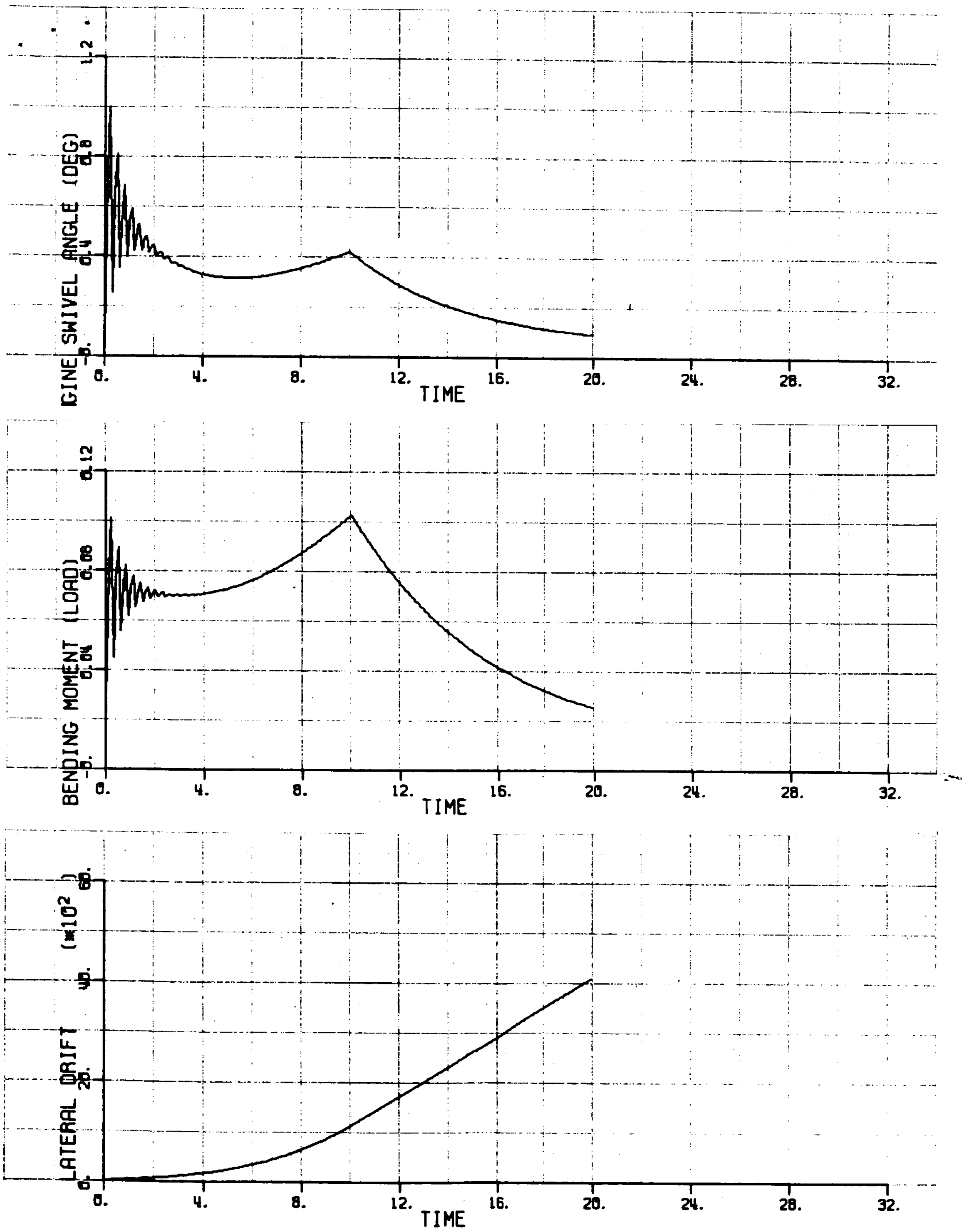


Figure 6-8. Linear control (continued).

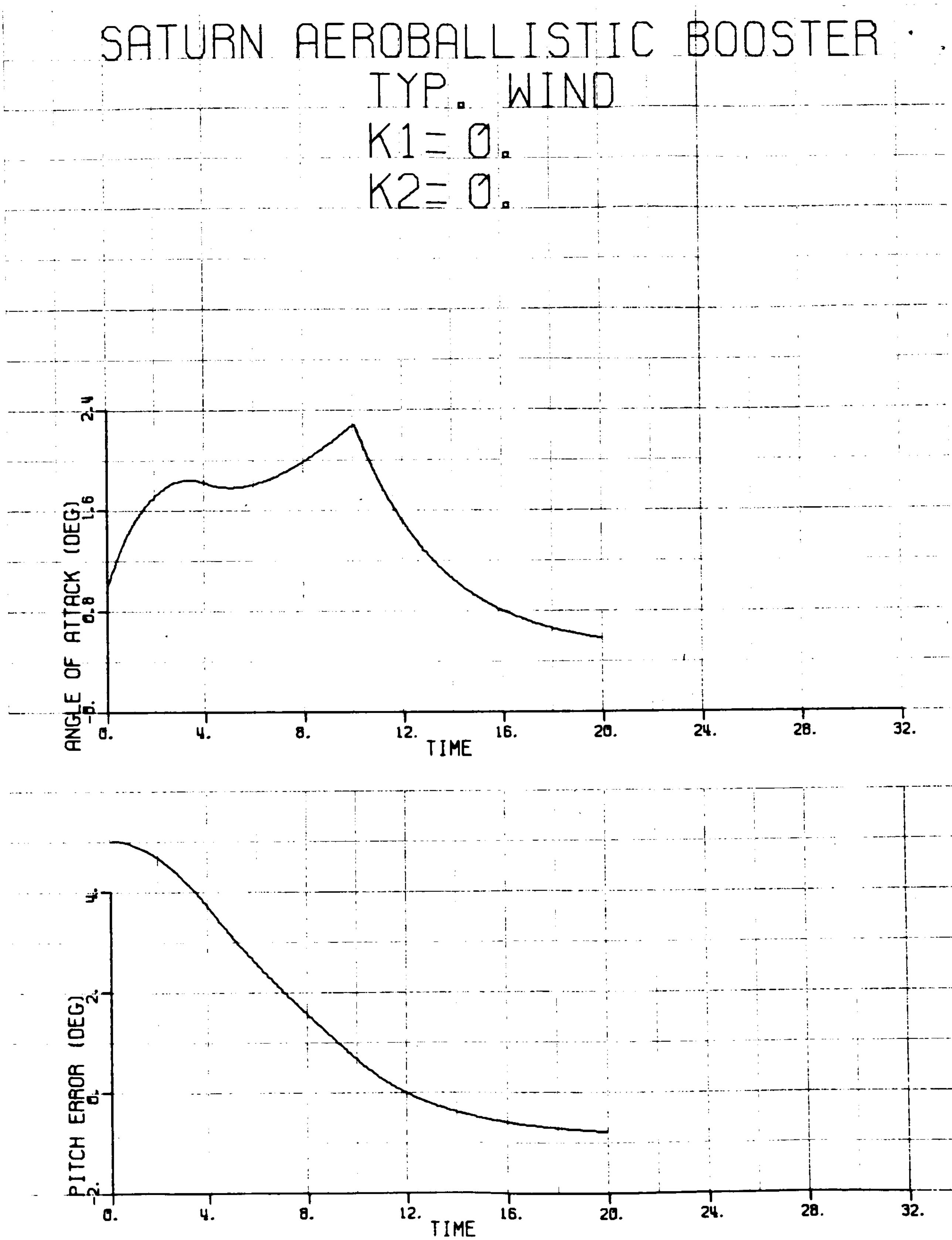


Figure 6-9. Bang-bang control.

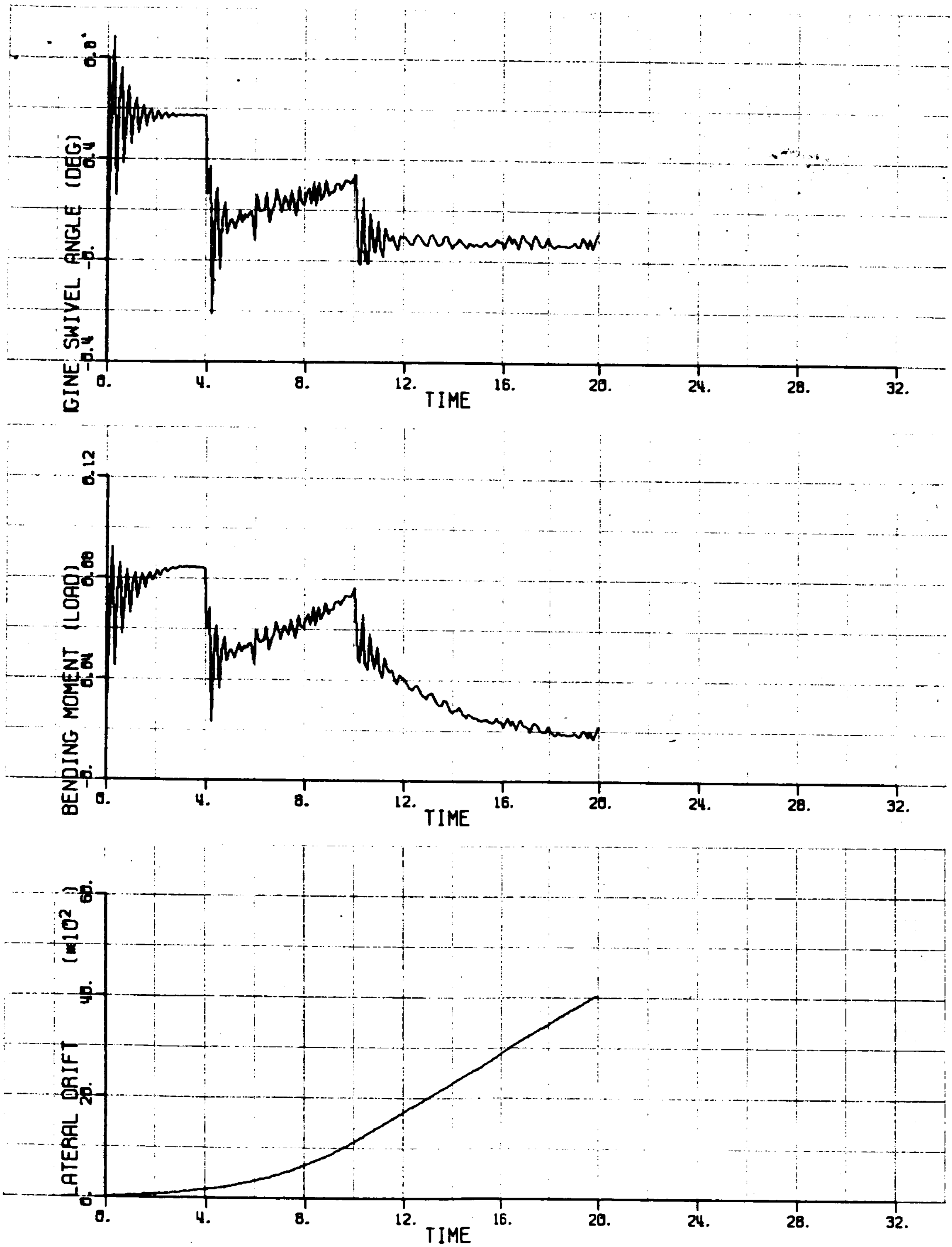


Figure 6-9. Bang-bang control (continued).

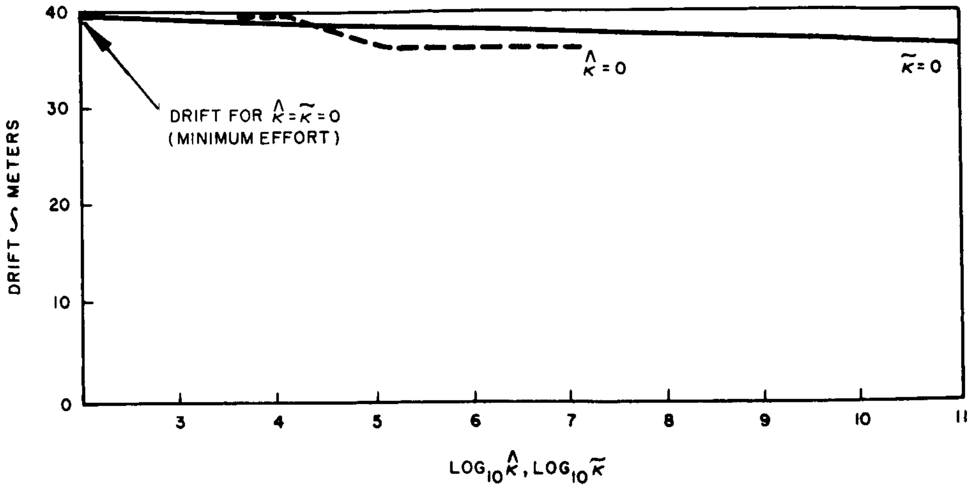


Figure 6-10. Drift after ten seconds for initial offset of $\alpha = 0.1°$, $\phi = 0.5°$.

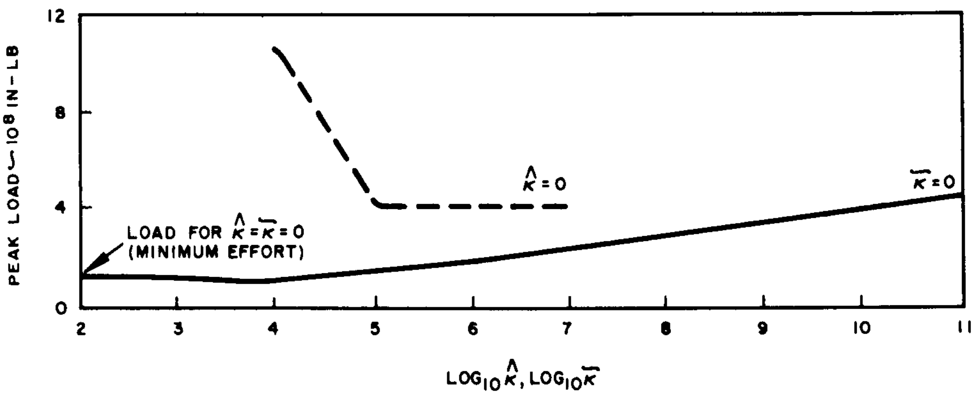


Figure 6-11. Peak load for initial offsets of $\alpha = 0.1°$, $\phi = 0.5°$, linear control.

a monotonically decreasing function $\tilde{\kappa}$, but is ultimately so. Since (6-5) shows that L is largely dependent upon β, the peak load must depend upon the type of control chosen. For linear control in the absence of wind the peak load is linearly related to the initial offsets, while the peak load for bang-bang control is much less sensitive to initial conditions. However, after the initial response transients died out, the "long-term" (5-10 seconds) behavior of the bending load was reasonably close for both controls. In the presence of wind, the situation is much more complicated (as Figures 6-2 through 6-9 show). The peak load is now a function of initial offset, control, and wind profile. Since the available simulation time was severely limited, there is insufficient data on which to base any realistic conclusions about the load behavior in the presence of wind.

The final series of flights consisted of a seven-dimensional model (including one bending mode) using the five-dimensional control vector discussed above for the case $\hat{\kappa} = \tilde{\kappa} = 0$, including the effect of the wind. The equations used were

$$\dot{\alpha} = -0.0322\alpha - 0.0194\phi + \dot{\phi} - 0.0211\beta + \dot{\alpha}_w \tag{6-11}$$

$$\ddot{\phi} = 0.0693\alpha - 0.474\beta \tag{6-12}$$

$$\ddot{\beta} = 0.762\alpha - 1760.5\beta - 3.36\dot{\beta} - 0.0357\dot{\zeta} + 17.5\psi \tag{6-13}$$

$$\ddot{\zeta} = 15.2\ddot{\beta} - 0.0036\zeta - 0.006\dot{\zeta} \; . \tag{6-14}$$

The performance of booster control system was not much affected by the inclusion of the bending term; the peak bending load was reduced somewhat, and the drift was unaffected.

7. DISCUSSION AND SUGGESTIONS

In Section 1 a 26-pole model of a large flexible vehicle was described. The model presented there can be used for extensions of the initial synthesis procedures described in Section 5, which were based on a 5-dimensional model, or for checking the stability and performance of closed loop systems for which the control law was derived for a lower dimensional model. It is felt that by adequately describing the dynamics of the vehicle, more reliable information about the stability of the actual vehicle may be obtained. In particular, the problem of blending sensor outputs so as to accurately identify the "state" of the system is intimately related to the existence of a complete model of the vehicle. This is an area of research which has recently received attention at Hughes Aircraft Company.

We propose to couple an accurate model of the vehicle with Kalman filtering and the filtering technique of Hughes Aircraft Company described in Section 3 to synthesize a system which would be immune to the noise introduced by the sensors, be insensitive to the higher bending modes of the vehicle, and minimize a given performance index. Hughes Aircraft Company has recently obtained results along these lines which are presently being evaluated.

The linear design procedure described in Section 5 may be extended to a higher dimensional model with little difficulty, but even for the existing procedure there are some areas of investigation which should be explored. First it should be determined how good the control law developed for a 5-dimensional model would perform when used for the 26-pole model. This has already been done for a 7-pole model. Also the possible improvement of performance gained by going from a 5-dimensional model to a 26-dimensional model should be explored to see if the added complexity is justified when the disturbance is the worst wind.

The nonlinear feedback law described in Section 4 seems to yield a minimax type of response when the system is subjected to an initial condition. This can be further checked by starting near the origin and flooding the state space to determine actually how good the performance is. It should be noted that the number of products of state variables increases factorially with the order of the model. It would, therefore, be worthwhile to investigate the performance of a high order system subjected to disturbances when the nonlinear control law has been designed for a low order plant. There is good reason to believe that the resulting system will be stable since feedback, in general, compensates for ignorance about the actual plant dynamics.

In the area of stability of closed loop system the results of Sections 3 and 4 indicate that the control laws derived there are globally asymptotically stable when there is no actuator saturation. For systems where there is actuator saturation the resulting systems are still stable in a well-defined, computable neighborhood or the origin. The results along these lines are presented in Appendices C and E. However, the stability results derived so far relate only to initial conditions and not to continuously acting disturbances. In order to determine the actual behavior of a working system it is necessary to obtain analytical results which will allow one to say exactly in what region the system is operating. This requires using the concept of "practical stability" as defined by Lasalle and Lefschetz to get an accurate assessment of the behavior of the system for "worst" input disturbances. In the case of linear systems bounds are easily obtained and, in fact, were presented in Hughes Aircraft Company's original technical proposal. However, in the case of nonlinear controlled systems or linearly controlled systems for which the control law was derived for a lower order model, these bounds are not easily obtained and further work is necessary to get an accurate assessment of behavior of the system. This would correspond to completing the study of the performance index

$$\min_{\psi \in \Psi} \|x\|$$

where one takes the maximum overall allowable disturbances.

Thus we conclude that this study has yielded some very useful control laws for linear models of the vehicle when one considers initial condition disturbances, but also there should be further study to assess their usefulness when applied to linear plants acted upon by "worst" disturbances.

APPENDIX A

LINEAR CANONICAL FORMS FOR CONTROLLABLE SYSTEMS

by

R. W. Bass and I. Gura
Space Systems Division,
Hughes Aircraft Company
under

Contract No. NAS8-11421

Space Systems Division

AEROSPACE GROUP

Hughes Aircraft Company . Culver City, California

APPENDIX A

LINEAR CANONICAL FORMS FOR CONTROLLABLE SYSTEMS

INTRODUCTION

In this paper four different coordinate systems are studied, namely

1) state variables (x)
2) phase coordinates (θ)
3) Lur'e coordinates (ξ)
4) generalized Lur'e coordinates (ϕ)

There are six nonsingular linear transformations, namely

$$\phi = T\theta$$

$$x = D\phi = DT\theta$$

$$\xi = V^*x = V^*D\phi = V^*DT\theta$$

that relate the four coordinate systems. In order to pass freely among these coordinate systems, including the inverse transformations, a total of twelve matrices must be utilized.

In particular numerical applications wherein the dimension n of the state space is large, it is desirable to avoid either inversion of n x n matrices, or complete spectral analyses of (nonsymmetric) matrices. The present analysis achieves this by explicit presentation in "closed form" of rational expressions for the elements of all twelve matrices.

It has been shown by Lur'e [1], Letov [2], and many others, that use of Lur'e coordinates facilitates explicit construction of Liapunov functions [3], hence facilitates study of stability of equilibrium in dynamical systems.

More recently it has been shown by Bass, Lewis, and Mendelson [4], [5], by Wonham and Johnson [6], [7], [8], by Kalman [9], and by Bass and Gura [10] that use of phase coordinates facilitates the application of frequency-domain concepts to various problems of system stabilization and optimization stated in time-domain concepts.

In this paper a system of generalized Lur'e coordinates is defined. Unlike the Lur'e coordinates, these variables are well-defined regardless of whether or not the system's "open-loop poles" (eigenvalues, characteristic roots) are distinct. Although many realistic engineering problems do not have multiple roots, many highly illuminating examples of modern control theory can be derived readily when multiple roots are permitted. Therefore the complete generality of applicability of this last-mentioned coordinate system is important for both exposition and research on advanced control problems.

The system to be studied is of the type

$$\dot{x} = Ax + a\psi_o$$

where

$$\dot{x} = Ax$$

governs the evolution in time of the uncontrolled system, where "a" is the actuator vector, and where the scalar $\psi_o = \psi_o(x)$ denotes the feedback control law. (In this paper the functional nature of ψ_o is irrelevant, hence unspecified.)

The characteristic polynomial of the uncontrolled system is defined by

$$\Delta(s) = \det(sI - A) = \sum_{i=0}^{n} \alpha_i s^i ,$$

which defines the coefficients $\alpha_o, \alpha_1, \cdots, \alpha_{n-1}, \alpha_n = 1$. Similarly, matrices $S_1, S_2, \cdots, S_n$ are defined either by

$$S_i = \sum_{j=i}^{n} \alpha_j A^{j-i} , \qquad (i = 1, 2, \cdots, n),$$

or by means of the resolvent equation

$$(sI - A)^{-1} = \sum_{i=1}^{n} \left\{ \frac{s^{i-1}}{\Delta(s)} \right\} S_i \ .$$

In numerical practice, use of the preceding definitions for the α_i and S_i is quite impossible for large values of n, since it would require n! multiplications. However, an efficient recursive algorithm stated below permits their computation in about n^4 multiplications.

The given system is called controllable [9] if the system of n simultaneous linear equations

$$a \cdot b = 0 \ , \quad Aa \cdot b = 0 \ , \quad \cdots , \quad A^{j-1} a \cdot b = 0 \ , \quad \cdots ,$$
$$A^{n-2} a \cdot b = 0 \ , \quad A^{n-1} a \cdot b = 1 \ ,$$

has a unique vector $b \neq 0$ for its solution. The vector b can be computed by Gaussian elimination. In general, computing b represents $(1/n)^{\text{th}}$ of the arithmetic labor required to invert an $n \times n$ matrix.

The vector b is important for several reasons. In particular, it is the normal vector at $x = 0$ to the time-optimal switching surface of the given control problem. In fact, it can be proved [11], [12] that the time-optimal regulator law has the form

$$\psi_o = \operatorname{sgn}[b \cdot x + \rho_o(x)] \ ,$$

where $\{\rho_o(x)/\|x\|\} \to 0$ as $\|x\| \to 0$; in fact for some $\epsilon_1 > 0$ there are positive constants μ_o, η_o such that

$$|\rho_o(x)| \leq \mu_o \|x\|^{1+\eta_o} \ , \qquad \eta_o > 0 \ , \qquad (\|x\| \leq \epsilon_1) \ .$$

Furthermore, if the phase variable θ_1 is defined by

$$\theta_1 = b \cdot x \ ,$$

then it will be shown below that the given system is equivalent to the scalar system of $n^{\underline{th}}$ order defined by

$$\Delta\left(\frac{d}{dt}\right)\theta_1 = \psi_o \ .$$

Passage from the phase variables θ_1, $\dot{\theta}_1, \cdots$, $d^{j-1}\theta_1/dt^{j-1}, \cdots$, $d^{n-1}\theta_1/dt^{n-1}$ to the state variables x_1, $x_2, \cdots$, x_n is facilitated by the result

$$x = \sum_{i=1}^{n} \left\{\frac{d^{i-1}\theta_1}{dt^{i-1}}\right\} S_i a$$

to be proved below.

Next, assume distinct roots, i.e. assume that the complex numbers λ_1, $\lambda_2, \cdots$, λ_n satisfy

$$\Delta(\lambda_i) = 0 \ , \qquad \Delta'(\lambda_i) \neq 0 \ , \qquad (i = 1, 2, \cdots, n) \ .$$

Define vectors v^i as suitably normalized eigenvectors of A^*, namely,

$$A^* v^i = \lambda_i v^i \ , \qquad v^i \cdot a = 1 \ , \qquad (i = 1, 2, \cdots, n) \ .$$

Then the <u>Lur'e coordinates</u> of x are given by

$$\xi_i = v^i \cdot x \ , \qquad (i = 1, 2, \cdots, n) \ ;$$

it is easy to see that these variables satisfy the system

$$\dot{\xi}_i = \lambda_i \xi_i + \psi_o \ , \qquad (i = 1, 2, \cdots, n) \ .$$

Furthermore, it will be proved that return from the variables ξ_i to the x_i is provided by the transformation

$$x = \sum_{i=1}^{n} \xi_i u^i ,$$

where the vectors u^i are defined as suitably normalized eigenvectors of A, namely

$$Au^i = \lambda_i u^i , \qquad u^1 + u^2 + \cdots + u^n = (1, 1, \cdots, 1)^* .$$

The preceding definitions of the u^i and v^i are adequate in principle but in practice are inconvenient. However, the correctly normalized u^i and v^i can be computed efficiently by the following closed form expressions:

$$u^i = \sum_{j=1}^{n} \left\{ \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \right\} S_j a , \qquad (i = 1, 2, \cdots, n) ,$$

$$v^i = \sum_{i=1}^{n} (\lambda_i)^{j-1} S_j^* b , \qquad (i = 1, 2, \cdots, n) .$$

A complete summary of results, in systematic tabular form, are given at the end of this appendix. All of these formulas are used in the authors' theory of integrals and isochrones [11] which allows explicit (local) solution in closed ("algebroid") form of both the time-optimal regulator problem [12] and the bang-bang control problem with quadratic performance index [13].

NATIONAL CONVENTIONS

a. Matrices are upper case letters.

b. Vectors are lower case unsubscripted or superscripted letters.

c. Scalars are subscripted lower case letters.

d. Exceptions to these rules are i, j, k, l, ν, n which are used as summation indices or scalars; s which is a complex scalar; $\Delta(s)$ which is a polynomial in s; and t which denotes time.

e. Asterisks used as superscripts (*) denote matrix transposition.

f. The i^{th} column of the identity matrix is represented by e^i.

g. The symbol $\triangleq$ denotes equality by definition.

ALGEBRAIC PRELIMINARIES

In general, the solution of the system of differential equations

$$\dot{x} = Ax + a\psi_o \tag{1}$$

involves the transition matrix e^{At}, whose Laplace transform is the resolvent matrix $(sI - A)^{-1}$ where I is the identity matrix and s is a scalar. It can be shown [4, 14] that this matrix is given by

$$(sI - A)^{-1} = \frac{\Gamma(s)}{\Delta(s)} \tag{2}$$

where

$$\Delta(s) = \det(sI - A) = \sum_{j=0}^{n} \alpha_j s^j \, , \qquad \Gamma(s) = \sum_{i=1}^{n} s^{i-1} S_i \, , \tag{3}$$

and the $S_1, S_2, \cdots, S_n$ and the $\alpha_o, \alpha_1, \cdots, \alpha_n$ are effectively computable by the recursion relations

$$\alpha_n = 1 \ , \qquad S_n = I \tag{4a}$$

$$\alpha_{n-j} = -\frac{1}{j} \operatorname{tr}(AS_{n-j+1}) \ , \qquad (j = 1, 2, \cdots, n) \ , \tag{4b}$$

$$S_{n-j} = \alpha_{n-j} I + AS_{n-j+1} \ , \qquad (j = 1, 2, \cdots, n) \ , \tag{4c}$$

The matrices S_i can be shown [4] to satisfy

$$S_{n-j} = \sum_{i=n-j}^{n} \alpha_i A^{i-n+j} \qquad (j = 1, 2, \cdots, n) \ . \tag{4d}$$

The theoretical definitions (3) and (4d) cannot be used to compute the α_i and S_i for large n, as they involve n! multiplications. However, the algorithm (4b-c) requires only about n^4 multiplications and has an intrinsic self-checking feature in that (by Cayley-Hamilton) $S_o = 0$.

The controllability criterion of Kalman [9] is fundamental to the present analysis and will be assumed henceforth. For the system (1) it can be expressed in determinantal form as

$$\det D \neq 0 \tag{5a}$$

where

$$D = (a, Aa, \cdots, A^{n-1}a) . \tag{5b}$$

<u>Theorem 1</u>

If the matrix L is defined implicitly by

$$L^{-1} \triangleq (S_1 a, S_2 a, \cdots, S_n a)^* \tag{6}$$

then

$$L \equiv \left[b, A^*b, (A^*)^2 b, \cdots, (A^*)^{n-1} b\right] \tag{7}$$

where the vector b is given by the solution (e. g., by Gaussian elimination) of the nonsingular system of linear equations

$$D^* b = e^n \tag{8}$$

Proof. If the above hypothesis is to be identically true, it must be shown that

$$[(S_1 a, S_2 a, \cdots, S_n a)^*]^{-1} e^i = (A^*)^{i-1} b \ , \qquad (i = 1, 2, \cdots, n) \tag{9a}$$

or, equivalently, that

$$e^i = (S_1 a, S_2 a, \cdots, S_n a)^* (A^*)^{i-1} b \ , \qquad (i = 1, 2, \cdots, n) \tag{9b}$$

is valid. In particular, the rows of (9b) can be written as

$$a^* S_j^* (A^*)^{i-1} b = a^* \sum_{\nu=j}^{n} \alpha_\nu (A^*)^{\nu-j+i-1} b = \delta_{ij} \ , \qquad (i, j = 1, 2, \cdots, n) \tag{10}$$

Now replace summation over ν by summation over k where $k = \nu + i - j$, and obtain

$$a^* \sum_{k=i}^{n+i-j} \alpha_{k+j-i} (A^*)^{k-1} b = \delta_{ij} \tag{11}$$

as the relationship to be established.

Consider first the case for which $j \geq i$. This implies that $1 \leq k \leq n$. Note that (8) can be written explicitly as

$$\delta_{kn} = a^* (A^*)^{k-1} b \ , \qquad (k = 1, 2, \cdots, n) \tag{12}$$

where δ_{kn} is the Kronecker delta. With this, the left side of (11) becomes $\Sigma_{k=1}^{n-j+i}\ \alpha_{k+j-i}\,\delta_{kn}$. The summand is zero except when k = n (which requires i = j) in which case the sum takes the value $\alpha_n = 1$. Hence (11) is true for $j \ge i$.

Returning to (11) when $j < i$, write the left side of that equation as

$$a^* \sum_{k=i}^{n} \alpha_{k+j-i}(A^*)^{k-1}b + a^* \sum_{k=n+1}^{n-j+i} \alpha_{k+j-i}(A^*)^{k-1}b \tag{13}$$

Now, by the same argument used above, the first summation in (13) yields the value α_{n+j-i}. On replacing k by $m = k + j - i$, the second sum becomes

$$a^*A^{i-j-1} \sum_{m=n+1+j-i}^{n} \alpha_m (A^*)^m b = -a^*A^{i-j-1} \sum_{m=0}^{n+j-i} \alpha_m (A^*)^m b \ , \tag{14}$$

where the latter result was obtained by use of the Cayley-Hamilton Theorem. (A matrix satisfies its own characteristic equation.) Now since $j < i$, (12) can be used (with m instead of k) and the second sum equals

$$-\sum_{m=0}^{n+j-i} \alpha_m a^*(A^*)^{m+i-j-1}b = -\sum_{m=0}^{n+j-i} \alpha_m \delta_{m+i-j,\, n} \tag{15}$$

This has the value zero except when $m + i - j = n$ in which case it becomes $-\alpha_{n+j-i}$. Combining this result with that following (13), it is seen that for $j < i$ the left side of (11) is zero. Thus relationship (11) has been proven and theorem must be valid.

Theorem 2

A more concise expression for the inverse of L is

$$(L^{-1})^* = DT \tag{16}$$

where

$$T = T^* \triangleq \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & 1 \\ \alpha_2 & \alpha_3 & \cdots & 1 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \alpha_{n-1} & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{17}$$

Proof. By inspection, the $i^{\underline{th}}$ column of T can be written as

$$t^i = \sum_{j=i}^{n} \alpha_j \, e^{j-i+1} \tag{18}$$

Now by definition

$$DT = (Dt^1, Dt^2, \cdots, Dt^n) \tag{19}$$

where

$$Dt^i = \sum_{j=i}^{n} \alpha_j D \, e^{j-i+1} = \sum_{j=i}^{n} \alpha_j (A)^{j-i} a \tag{20}$$

But by (4), the definition of S_i, $Dt^i = S_i a$. Then applying (6) yields

$$DT = (L^{-1})^* = (S_1 a, S_2 a, \cdots, S_n a) \tag{21}$$

as desired.

Theorem 3

A pair of explicit expressions for the inverse of D is

$$D^{-1} \triangleq (a, Aa, \cdots, A^{n-1}a)^{-1} \equiv TL^* \tag{22a}$$

$$D^{-1} \equiv (S_1^*b, S_2^*b, \cdots, S_n^*b)^* \tag{22b}$$

Proof. Consider the matrix

$$LT^* = LT = (Lt^1, Lt^2, \cdots, Lt^n) \tag{23}$$

By (18) and the definition of L,

$$Lt^i = \sum_{j=i}^{n} [b, A^*b, \cdots (A^*)^{n-1}b]\alpha_j \, e^{j-i+1}$$

$$= \sum_{j=i}^{n} \alpha_j (A^*)^{j-i} b \;, \qquad (i = 1, 2, \cdots, n) \tag{24}$$

Applying (4d) it is seen that $Lt^i = S_i^*b$. Thus,

$$LT^* = (S_1^*b, S_2^*b, \cdots, S_n^*b) \tag{25}$$

Now by Theorem 2, $D^{-1} = [(L^{-1})^*T^{-1}]^{-1} = TL^*$, or $LT^* = (D^{-1})^*$ so that by (25)

$$D^{-1} = (S_1^*b, S_2^*b, \cdots, S_n^*b)^* \tag{26}$$

as required.

Theorem 4

An explicit expression for the inverse of T is

$$T^{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & \beta_1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 1 & \cdots & \beta_{n-3} & \beta_{n-2} \\ 1 & \beta_1 & \cdots & \beta_{n-2} & \beta_{n-1} \end{bmatrix} \tag{27}$$

where the β's are given by the following recursion formula

$$\beta_o = 1 \ , \tag{28a}$$

$$\beta_\nu = -\sum_{j=0}^{\nu-1} \alpha_{j+n-\nu} \beta_j \ , \qquad (\nu = 1, 2, \cdots, n-1) \tag{28b}$$

Proof. The proof of this theorem consists of two parts. The first part introduces the appropriate set of quantities (β_i) which obey (28). The second part shows that T^{-1} is given by the matrix displayed in (27).

Part A. Define the quantities $\beta_j (j = 1, 2, \cdots)$ by the Laurent series

$$\frac{1}{\Delta(s)} = \sum_{j=0}^{\infty} \frac{\beta_j}{s^{n+j}} \ , \qquad (|s| > \max |s_i|) \tag{29}$$

where the s_i are the roots of $\Delta(s)$. Then

$$1 = \left(\sum_{i=0}^{n} \alpha_i s^i\right)\left(\sum_{j=0}^{\infty} \beta_j s^{-(n+j)}\right) \tag{30}$$

Replace j by use of the definition $\nu = j + n - i$, obtaining

$$1 = \sum_{i=0}^{n} \sum_{\nu=n-i}^{\infty} \alpha_i \beta_{i+\nu-n} s^{-\nu} \tag{31}$$

Now interchange the order of summation by observing that $0 \le n - i \le \nu \le \infty$ and $0 \le i \le n$ imply that $0 \le \nu \le \infty$ and $\max(n - \nu, 0) \le i \le n$. Thus

$$1 = \sum_{\nu=0}^{\infty} \left\{ \sum_{i=\max(n-\nu,\,0)}^{n} \alpha_i \beta_{i+\nu-n} \right\} s^{-\nu} \tag{32}$$

Note that the very first term on the right side of (32) is the only constant in the series. Thus for (32) to be valid for all $|s| \ge \max |s_i|$ that term must be equal to unity and the remaining terms must all be zero. Then

$$\alpha_n \beta_o = 1 \tag{33a}$$

$$\sum_{i=n-\nu}^{n} \alpha_i \beta_{i+\nu-n} = 0 \ , \qquad (\nu = 1, 2, \cdots, n) \tag{33b}$$

$$\sum_{i=0}^{n} \alpha_i \beta_{i+\nu-n} = 0 \ , \qquad (\nu = n+1, n+2, n+3, \cdots) \ , \tag{33c}$$

or equivalently, $\beta_o = 1$,

$$\beta_\nu = - \sum_{i=n-\nu}^{n-1} \alpha_i \beta_{i+\nu-n} = - \sum_{j=0}^{\nu-1} \alpha_{j+n-\nu} \beta_j \ , \qquad (\nu = 1, 2, \cdots, n) \ , \tag{34a}$$

where $j = i + \nu - n$, and, similarly,

$$\beta_{k+n} = - \sum_{j=k}^{k+n-1} \alpha_{j-k} \beta_j \ , \qquad (k = 1, 2, \cdots,) \ , \tag{34b}$$

where $k = \nu - n$.

Part B. It will be shown that $TT^{-1} = I$, where T^{-1} is defined by (27). By inspection, the $j^{\underline{th}}$ column of T^{-1} is given by

$$\tau^{j} = \sum_{k=0}^{j-1} \beta_k \, e^{n+k-j+1} \tag{35}$$

Then, using (18), the i-$j^{\underline{th}}$ element of $TT^{-1} = T^{*}T^{-1} = (T^{*}\tau^{1}, \cdots, T^{*}\tau^{n})$ is

$$t^{i} \cdot \tau^{j} = \sum_{\ell=i}^{n} \sum_{k=0}^{j-i} \alpha_{\ell} \beta_k \, \delta_{\ell-i+1,\, n+k-j+1} \tag{36}$$

The non-zero terms of this expression occur only when $\ell - i + 1 = n + k - j + 1$ or when $\ell = n + k - j + i$. However, $i \leq \ell \leq n$ and $0 \leq k \leq j - 1$ must also be satisfied. This implies that $i \leq n + k - j + i \leq n$ or that $0 \leq k \leq j - i$. Then (36) becomes

$$t^{i} \cdot \tau^{j} = \sum_{k=0}^{j-i} \alpha_{n+k-j+i} \beta_k \tag{37}$$

For $j = i$ this reduces to unity. For $j \neq i$ let $\nu = j - i$ and, using (34a), obtain

$$t^{i} \cdot \tau^{j} = \sum_{k=0}^{\nu} \alpha_{n+k-\nu} \beta_k = -\beta_{\nu} + \beta_{\nu} = 0 \tag{38}$$

and the theorem is proven.

PHASE VARIABLES (θ)

Taking the scalar product of $(A^{*})^{k-1}b$, $(k = 1, 2, \cdots, n)$, with the system (1) results in

$$\left[(A^{*})^{k-1} b \cdot \frac{dx}{dt}\right] = (A^{*})^{k-1} b \cdot Ax + (A^{*})^{k-1} b \cdot a\psi_o. \tag{39}$$

Applying (12) gives

$$\left[(A^*)^{k-1}b \cdot \frac{dx}{dt}\right] = (A^*)^k b \cdot x + \delta_{kn} \psi_o . \tag{40}$$

Now define a new variable

$$\theta_1 = b \cdot x \tag{41}$$

where b satisfies (8). Then for $k = 1$, (40) becomes

$$b \cdot \frac{dx}{dt} = \frac{d\theta_1}{dt} = A^* b \cdot x \tag{42}$$

Differentiating this expression with respect to time and using (40) for $k = 2$ gives

$$\frac{d^2\theta_1}{dt^2} = A^* b \cdot \frac{dx}{dt} = (A^*)^2 b \cdot x \tag{43}$$

Continuing in this manner obtain

$$\frac{d^{i-1}\theta_1}{dt^{i-1}} = (A^*)^{i-1} b \cdot x \qquad (i = 1, 2, \cdots, n) \tag{44a}$$

and

$$\frac{d^n\theta_1}{dt^n} = (A^*)^n b \cdot x + \psi_o . \tag{44b}$$

Then

$$\sum_{j=0}^{n} \alpha_j \frac{d^j\theta_1}{dt^j} = [\alpha_o I + \alpha_1 A^* + \cdots + \alpha_n (A^*)^n] b \cdot x + \psi_o . \tag{45}$$

Now by the Cayley-Hamilton Theorem $\Delta(A^*) = 0$, whence

$$\sum_{j=0}^{n} \alpha_j \frac{d^j\theta_1}{dt^j} = \Delta(d/dt)\theta_1 = \psi_o \tag{46}$$

Upon defining the state variables $\theta_1, \theta_2, \cdots, \theta_n$ by

$$\theta_i = \frac{d^{i-1}\theta_1}{dt^{i-1}}, \qquad (i = 1, 2, \cdots, n), \tag{47}$$

the $n^{\underline{th}}$ order scalar differential equation (46) can be expressed as the first order matrix system

$$\dot{\theta} = C\theta + e^n \psi_o \tag{48a}$$

where

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdot \\ \cdot \\ \cdot \\ \theta_n \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-2} & -\alpha_{n-1} \end{bmatrix} \tag{48b}$$

To find the transformation matrix between the x and the θ coordinates, note that Equation (44a) can be expressed as

$$\theta_i = (A^*)^{i-1} b \cdot x, \qquad (i = 1, 2, \cdots, n), \tag{49a}$$

or

$$\theta = [b, A^* b, \cdots, (A^*)^{n-1} b]\ x = L^* x \tag{49b}$$

Note that applying this directly to (1) and comparing the result with (48) shows that

$$C = \overset{*}{L} A (L^*)^{-1} \tag{50}$$

By Theorems 1 and 2 the inverse of (49b) can be established directly. Thus

$$x = (L^*)^{-1}\theta = (S_1 a, S_2 a, \cdots, S_n a)\theta = \sum_{i=1}^{n} \theta_i S_i a \tag{51a}$$

or

$$x = DT\theta$$

"GENERALIZED" LUR'E VARIABLES (ϕ)

(The reason for this name will become clear in a later section.)

<u>Relations Between x and ϕ</u>

Let

$$\phi \triangleq D^{-1} x \tag{52}$$

Then (1) becomes

$$\phi = (D^{-1} A D)\phi + D^{-1} a \psi \tag{53}$$

Consider now the matrix product

$$DC^* = (a, Aa, \cdots, A^{n-1}a)\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \cdots & 0 & -\alpha_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{bmatrix}$$

$$= (Aa, A^2a, \cdots, -\sum_{i=0}^{n-1} \alpha_i A^i a) \ . \tag{54}$$

Applying the Cayley-Hamilton Theorem, the last column of (54) becomes $A^n a$ whence

$$DC^* = AD \tag{55a}$$

or

$$D^{-1}AD = C^* \tag{55b}$$

Note also that, by Theorem 3,

$$D^{-1}a = (S_1^*b, S_2^*b, \cdots, S_n^*b)^* a \tag{56a}$$

or, using Equation (10),

$$D^{-1}a = \begin{bmatrix} a \cdot S_1 b \\ a \cdot S_2 b \\ \cdot \\ \cdot \\ \cdot \\ a \cdot S_n b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = e^1 \tag{56b}$$

Thus (53) can be expressed as

$$\dot{\phi} = C^{*}\phi + e^{1}\psi_{o} . \tag{57}$$

The forward and reverse transformation relations can be expressed explicitly as follows. By (52) and Theorem 3,

$$\phi = D^{-1}x = TL^{*}x = (S_{1}^{*}b, S_{2}^{*}b, \cdots, S_{n}^{*}b)^{*}x , \tag{58a}$$

or

$$\phi_{i} = (S_{i}^{*}b) \cdot x , \qquad (i = 1, 2, \cdots, n) . \tag{58b}$$

Also

$$x = D\phi = (a, Aa, \cdots, A^{n-1}a)\phi = \sum_{i=1}^{n} \phi_{i}A^{i-1}a \tag{59}$$

<u>Relations Between θ and ϕ</u>

Previously [(58a) and (51b)] it has been established that

$$\phi = D^{-1}x , \qquad x = DT\theta . \tag{60}$$

Consequently,

$$\phi = T\theta \tag{61}$$

In particular, using (18)

$$\phi = (t^{1}, t^{2}, \cdots, t^{n})\theta = \sum_{i=1}^{n} t^{i}\theta_{i} = \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{j} e^{j-i+1}\theta_{i} , \tag{62}$$

and so

$$\phi_\nu = \phi \cdot e^\nu = \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_j \theta_i \delta_{\nu, j-i+1} , \qquad (\nu = 1, 2, \cdots, n) . \tag{63}$$

Non-zero terms occur in (63) only when $\nu = j - i + 1$ or when $j = \nu + i - 1$. Combining this with the constraints $1 \le i \le n$ and $i \le j \le n$, j can be replaced by $\nu + i - 1$ only if $1 \le i \le n - \nu + 1$. Then

$$\phi_\nu = \sum_{i=1}^{n-\nu+1} \alpha_{\nu+i-1} \theta_i \tag{64}$$

whence, setting $\ell = \nu + i - 1$

$$\phi_\nu = \sum_{\ell=\nu}^{n} \alpha_\ell \theta_{\ell-\nu+1} , \qquad (\nu = 1, 2, \cdots, n) , \tag{65a}$$

$$\phi_n = \theta_1 \tag{65b}$$

The inverse transformation can be established in a similar manner. Employing (35),

$$\theta_\nu = T^{-1} \phi \cdot e^\nu = \sum_{i=1}^{n} \tau^i \phi_i \cdot e^\nu = \sum_{i=1}^{n} \sum_{k=0}^{i-1} \beta_k \phi_i \, e^{n+k-i+1} \cdot e^\nu$$

$$= \sum_{i=1}^{n} \sum_{k=0}^{i-1} \beta_k \phi_i \delta_{\nu, n+k-i+1} \tag{66}$$

This expression can be simplified to

$$\theta_\nu = \sum_{i=n-\nu+1}^{n} \beta_{\nu-n+i-1} \phi_i \tag{67}$$

by considerations similar to those used after (63). Finally, if summation over i is replaced by summation over $\ell = \nu - n + i - 1$, there results

$$\theta_\nu = \sum_{\ell=0}^{\nu-1} \beta_\ell \phi_{\ell+n-\nu+1} , \qquad (\nu = 1, 2, \cdots, n) , \tag{68a}$$

$$\theta_1 = \phi_n \tag{68b}$$

LUR'E COORDINATES (ξ)

Relations Between ξ and ϕ

By inspection of Equations (54) and (57), the system (1) is precisely equivalent to

$$\dot{\phi}_1 = -\alpha_o \phi_n + \psi_o , \tag{69a}$$

$$\dot{\phi}_2 = \phi_1 - \alpha_1 \phi_n , \tag{69b}$$

$$\dot{\phi}_j = \phi_{j-1} - \alpha_{j-1} \phi_n , \quad (j = 2, 3, \cdots, n). \tag{69c}$$

Now consider the ϕ coordinates for a system with distinct complex eigenvalues λ_i, $(i = 1, 2, \cdots, n)$. Multiply the $j^{\underline{th}}$ equation in (69) by λ_i^{j-1} and sum to obtain

$$\sum_{j=1}^{n} \lambda_i^{j-1} \phi_j = \sum_{j=1}^{n-1} \lambda_i^j \phi_j - \sum_{j=0}^{n-1} \alpha_j \lambda_i^j \phi_n + \psi_o , \qquad (i = 1, 2, \cdots, n) \tag{70}$$

Now since

$$\Delta(\lambda_i) = 0 , \qquad -\sum_{j=0}^{n-1} \alpha_j \lambda_i^j = \alpha_n \lambda^n ,$$

and (70) reduces to

$$\sum_{j=1}^{n} \lambda_i^{j-1} \phi_j = \sum_{j=1}^{n} \lambda_i^{j} \phi_j + \psi_o, \qquad (i = 1, 2, \cdots, n) \tag{71}$$

Define

$$\xi_i \triangleq \sum_{j=1}^{n} \lambda_i^{j-1} \phi_j \tag{72}$$

as the $i^{\underline{th}}$ component of an n-vector ξ. Then (71) becomes

$$\dot{\xi}_i = \lambda_i \xi_i + \psi_o , \qquad (i = 1, 2, \cdots, n) \tag{73a}$$

or, in vector form,

$$\dot{\xi} = \Lambda \xi + u^o \psi_o \tag{73b}$$

where

$$\Lambda = (\lambda_1 e^1, \lambda_2 e^2, \cdots, \lambda_n e^n) , \qquad u^o = (1, 1, \cdots, 1)^* \tag{73c}$$

The transformation (72) between ξ and ϕ can be expressed in matrix form by the equation

$$\xi = Z^* \phi \tag{74}$$

where $Z = (z^1, z^2, \cdots, z^n)$ and where

$$z^i = \begin{bmatrix} 1 \\ \lambda_i \\ \lambda_i^2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_i^{n-1} \end{bmatrix} = \sum_{k=1}^{n} (\lambda_i)^{k-1} e^k , \qquad (i = 1, 2, \cdots, n) . \tag{75}$$

To find $(Z^*)^{-1}$ consider the following. The identity

$$Cz^i = \lambda_i z^i , \qquad (i = 1, 2, \cdots, n) \tag{76}$$

can be verified by inspection of (48b). Now by (55b), Theorem 3, and (50),

$$T^{-1}C^*T = T^{-1}D^{-1}ADT = L^*A(L^{-1})^* = C . \tag{77}$$

Hence

$$T^{-1}C^*Tz^i = \lambda_i z^i \tag{78}$$

or

$$C^*Tz^i = \lambda_i Tz^i . \tag{79}$$

If the λ_i, $(i = 1, 2, \cdots, n)$, are distinct, then $\Delta'(\lambda_i) = [d(\Delta(s))/ds]_{\lambda_i} \neq 0$ and so

$$C^* \frac{Tz^i}{\Delta'(\lambda_i)} = \lambda_i \frac{Tz^i}{\Delta'(\lambda_i)} . \tag{80}$$

Now define the vectors

$$w^i = Tz^i/\Delta'(\lambda_i) , \qquad (i = 1, 2, \cdots, n) . \tag{81a}$$

Then from (80),

$$C^*w^i = \lambda_i w^i , \qquad (i = 1, 2, \cdots, n) . \tag{81b}$$

Using (76) and (81) it is clear that

$$w^j \cdot Cz^i = \lambda_i w^j \cdot z^i , \qquad (i, j = 1, 2, \cdots, n) , \tag{82a}$$

and

$$z^i \cdot C^*w^j = \lambda_j z^i \cdot w^j , \qquad (i, j = 1, 2, \cdots, n) . \tag{82b}$$

Hence

$$\lambda_i(z^i \cdot w^j) \equiv \lambda_j(z^i \cdot w^j) \tag{83}$$

which implies that

$$z^i \cdot w^j = 0 \ , \qquad i \neq j \ . \tag{84}$$

For $i = j$, note that

$$z^i \cdot w^i = z^i \cdot \frac{Tz^i}{\Delta'(\lambda_i)} \tag{85}$$

By (75) and (18)

$$z^i \cdot Tz^i = \sum_{k=1}^{n} (\lambda_i)^{k-1} z^i \cdot t^k = \sum_{k=1}^{n} (\lambda_i)^{k-1} \sum_{\ell=k}^{n} \alpha_\ell \, z^i \cdot e^{\ell-k+1} \ . \tag{86}$$

Hence

$$z^i \cdot w^i = \sum_{k=1}^{n} \frac{(\lambda_i)^{k-1}}{\Delta'(\lambda_i)} \sum_{\ell=k}^{n} \alpha_\ell \lambda_i^{\ell-k} = \sum_{k=1}^{n} \sum_{\ell=k}^{n} \frac{\alpha_\ell(\lambda_i)^{\ell-1}}{\Delta'(\lambda_i)} \tag{87}$$

To reverse the order of summation in the last expression note that $1 \leq k \leq \ell \leq n$ implies $1 \leq \ell \leq n$ and $1 \leq k \leq \ell$. Thus (87) becomes, for $(i = 1, 2, \cdots, n)$,

$$z^i \cdot w^i = \sum_{\ell=1}^{n} \sum_{k=1}^{\ell} \frac{\alpha_\ell(\lambda_i)^{\ell-1}}{\Delta'(\lambda_i)} = \sum_{\ell=1}^{n} \frac{\ell\alpha_\ell(\lambda_i)^{\ell-1}}{\Delta'(\lambda_i)} = 1 \ . \tag{88}$$

Combining (84) and (88), there results $w^i \cdot z^i = \delta_{ij}$ or equivalently

$$(w^1, w^2, \cdots, w^n)^*(z^1, z^2, \cdots, z^n) = I \ . \tag{89}$$

If $W \triangleq (w^1, w^2, \cdots, w^n)$, then (89) becomes

$$W = (Z^*)^{-1} \tag{90}$$

Hence (74) implies

$$\phi = W\xi \ , \tag{91}$$

To express this relationship more explicitly note that, as in (86),

$$\phi = W\xi = \sum_{i=1}^{n} w^i \xi_i = \sum_{i=1}^{n} \frac{Tz^i}{\Delta'(\lambda_i)} \xi_i$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{(\lambda_i)^{k-1}}{\Delta'(\lambda_i)} \sum_{\ell=k}^{n} \alpha_\ell \, e^{\ell-k+1} \xi_i \ , \tag{92}$$

or

$$\phi_j = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{(\lambda_i)^{k-1}}{\Delta'(\lambda_i)} \sum_{\ell=k}^{n} \alpha_\ell \, \xi_i \, \delta_{j,\ell-k+1} \tag{93}$$

The summations are trivial except when $\ell = j + k - 1$. Combining this with the constraints $k \le \ell \le n$, $1 \le k \le n$, (93) reduces to

$$\phi_j = \sum_{i=1}^{n} \sum_{k=1}^{n-j+1} \frac{(\lambda_i)^{k-1}}{\Delta'(\lambda_i)} \alpha_{j+k-1} \xi_i \ , \tag{94}$$

or, setting $\nu = j + k - 1$,

$$\phi_j = \sum_{i=1}^{n} \left\{ \sum_{\nu=j}^{n} \frac{(\lambda_i)^{\nu-j}}{\Delta'(\lambda_i)} \alpha_\nu \right\} \xi_i \ , \qquad (j = 1, 2, \cdots, n) \ . \tag{95}$$

Relations Between ξ and θ

By (91) and (61) it is obvious that

$$\theta = T^{-1} W \xi \tag{96}$$

In particular, from (92)

$$\theta = \sum_{i=1}^{n} \frac{z^i}{\Delta'(\lambda_i)} \xi_i = \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{(\lambda_i)^{k-1}}{\Delta'(\lambda_i)} e^k \xi_i \ , \tag{97}$$

or

$$\theta_j = \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \xi_i \ , \qquad (j = 1, 2, \cdots, n) \ . \tag{98}$$

Similarly, the inverse transformation is easily established from (74) and (61) to be

$$\xi = Z^* T \theta \ . \tag{99}$$

Hence, proceeding as usual,

$$\xi_i = \sum_{j=1}^{n} z^i \cdot t^j \theta_j = \sum_{j=1}^{n} \sum_{k=j}^{n} \alpha_k \theta_j z^i \cdot e^{k-j+1}$$

$$= \sum_{j=1}^{n} \left\{ \sum_{k=j}^{n} \alpha_k \lambda_i^{k-j} \right\} \theta_j \ , \qquad (i = 1, 2, \cdots, n) \ . \tag{100}$$

Relations Between ξ and x

The basic relationship between ξ and x can be found immediately by applying (58b) to (72). Thus

$$\xi_i = \sum_{j=1}^{n} \lambda_i^{j-1} S_j^* b \cdot x \tag{101}$$

Now define $V \triangleq (v^1, v^2, \cdots, v^n)$, where

$$v^i = \sum_{j=1}^{n} \lambda_i^{j-1} S_j^* b \ , \qquad (i = 1, 2, \cdots, n) \tag{102}$$

Then

$$\xi_i = v^i \cdot x \ , \qquad (i = 1, 2, \cdots, n) \tag{103a}$$

or

$$\xi = V^* x \tag{103b}$$

Alternatively, combining (58a) and (74) gives, by Theorem 3,

$$\xi = Z^* T L^* x \ , \tag{104}$$

so that

$$V^* = Z^* T L^* \tag{105}$$

must be valid. By Theorem 2 and (90)

$$(V^*)^{-1} = (L^*)^{-1} T^{-1} (Z^*)^{-1} = DW \tag{106}$$

For convenience define

$$U \triangleq (u^1, u^2, \cdots, u^n) = DW$$

where, as in (92),

$$u^i = Dw^i = \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \sum_{k=j}^{n} \alpha_k A^{k-j} a = \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} S_i a \ . \tag{107}$$

Then

$$x = DW\xi = U\xi = \sum_{i=1}^{n} u^i \xi_i \ . \tag{108}$$

Extensions and Generalizations

The identity

$$\Delta(\eta) - \Delta(\mu) \equiv (\eta - \mu) \sum_{i=1}^{n} \eta^{i-1} \sum_{j=i}^{n} \alpha_j \mu^{j-i} \tag{109}$$

can easily be verified by equating coefficients of like powers of η and μ where these quantities obey the commutative and distributive laws of algebra. With no loss of generality, η can be identified with sI and μ with the matrix A. Then

$$\Delta(s)I - \Delta(A) = (sI - A) \sum_{i=1}^{n} s^{i-1} \sum_{j=i}^{n} \alpha_j A^{j-1} \tag{110}$$

and, by the Cayley-Hamilton Theorem and the definition of $\Gamma(s)$,

$$\Delta(s)I = (sI - A)\Gamma(s) \tag{111}$$

Indeed, (2) can be found directly from this relationship whenever $(sI - A)^{-1}$ exists. By multiplying (111) on the right by the vector a it is also clear that

$$\Delta(s)a = s\Gamma(s)a - A\Gamma(s)a \tag{112}$$

Before proceeding, define the vector u(s) by

$$u(s) \triangleq \frac{\Gamma(s)a}{\hat{\Delta}(s)} \tag{113}$$

where

$$\hat{\Delta}(s) = \begin{cases} \Delta(s) \text{ for } \Delta(s) \neq 0 \\ \Delta'(\lambda_i) \text{ for } \Delta(s) = \Delta(\lambda_i) = 0 \quad \text{and} \quad \lambda_i \neq \lambda_j\,, \quad (i, j = 1, 2, \cdots, n)\,. \end{cases}$$

Explicitly,

$$u(s) = \sum_{j=1}^{n} \sum_{k=j}^{n} \frac{s^{j-1}}{\hat{\Delta}(s)} \alpha_k A^{k-j} a = \sum_{k=1}^{n} \sum_{j=1}^{k} \frac{s^{j-1}}{\hat{\Delta}(s)} \alpha_k A^{k-j} a\,. \tag{114}$$

Now let $\ell = k - j + 1$ and replace j to obtain

$$u(s) = \sum_{k=1}^{n} \sum_{\ell=1}^{k} \alpha_k \frac{s^{k-\ell}}{\hat{\Delta}(s)} A^{\ell-1} a \tag{115}$$

Taking the scalar product of u(s) with the vector b and applying (12) it is clear that

$$u(s) \cdot b = \sum_{k=1}^{n} \sum_{\ell=1}^{k} \alpha_k \frac{s^{k-\ell}}{\hat{\Delta}(s)} \delta_{\ell n} = \frac{1}{\hat{\Delta}(s)} \tag{116}$$

Returning to (112), note that u(s) satisfies

$$\Delta(s)a + A\tilde{u}(s) = s\tilde{u}(s) \ , \qquad \Delta(s) \neq 0 \tag{117a}$$

where

$$\tilde{u}(s) \triangleq \hat{\Delta}(s)u(s) = \Gamma(s)a \tag{117b}$$

and so, dividing by $\Delta'(\lambda_i)$ and setting $s = \lambda_i$, there results

$$Au(\lambda_i) = \lambda_i u(\lambda_i) \ , \qquad \Delta(\lambda_i) = 0 \ , \qquad \lambda_i \neq \lambda_j (i, j = 1, 2, \cdots, n) \tag{117c}$$

$$u(\lambda_i) \cdot b = \frac{1}{\Delta'(\lambda_i)} \tag{117d}$$

In the latter case, the $u(\lambda_i)$ reduce exactly to the u^i defined in (107). Thus the columns of U are merely the eigenvectors of A, normalized according to (117d). Consider (109) again with η as sI and A^* as μ. As before, it can be shown that

$$\Delta(s)I = s\Gamma^*(s) - A^*\Gamma(s) \tag{118a}$$

or

$$\Delta(s)b + A^*\Gamma^*(s)b = s\Gamma^*(s) \tag{118b}$$

Define

$$v(s) \triangleq \Gamma^*(s)b \tag{119a}$$

or, equivalently,

$$v(s) = \sum_{j=1}^{n} s^{j-1} S_j^* b \ . \tag{119b}$$

Proceeding in a manner analogous to that followed in Equations (114) - (116), it is clear that

$$v(s) \cdot a = 1 \tag{120}$$

Also, by (118b)

$$\Delta(s)b + A^* v(s) = sv(s) \tag{121}$$

is always satisfied. When $\Delta(s) = \Delta(\lambda_i) = 0$, $(i = 1, 2, \cdots, n)$, (121) becomes

$$A^* v(\lambda_i) = \lambda_i v(\lambda_i) \ , \qquad (i = 1, 2, \cdots, n) \ , \tag{122a}$$

$$v(\lambda_i) \cdot a = 1 \qquad (i = 1, 2, \cdots, n) \ . \tag{122b}$$

By comparing (119b) and (122) with (102), it is obvious that $v(\lambda_i)$ is identical to v^i, $(i = 1, 2, \cdots, n)$, and that these vectors are the eigenvectors of A^* normalized according to (122b).

Note that (103a) can now be generalized, using (119b) and (58b), to

$$\xi_o(s) = v(s) \cdot x = \sum_{i=1}^{n} s^{i-1} \phi_i \ . \tag{123}$$

Then, taking the scalar product of v(s) with the system (1) and applying (120) and (121) it is found that

$$\begin{aligned} v(s) \cdot \dot{x} &= v(s) \cdot (Ax) + v(s) \cdot a\psi_o \\ &= x^* A^* v(s) + \psi_o \\ &= x^*(sv(s) - \Delta(s)b) + \psi_o \end{aligned} \tag{124}$$

Now using (123) and (49a), the above becomes

$$\dot{\xi}_o(s) = s\xi_o(s) - \Delta(s)\theta_1 + \psi_o , \qquad \theta_1 = b \cdot x = \phi_n \tag{125}$$

This can be considered a generalization of the Lur'e canonical form. In fact, when the eigenvalues of A are distinct,

$$\xi_i = \xi_o(\lambda_i) , \qquad (i = 1, 2, \cdots, n) , \tag{126}$$

and, setting $s = \lambda_i$ in (125), the Lur'e form (73a) is recovered. On the other hand, whether or not the λ_i are distinct, the identity (125), which in form is highly reminiscent of the Lur'e form, can be regarded as the collection of n differential equations obtained by equating like powers of s on the right and left hand sides. However, on inserting (123) into (125) and comparing coefficients, the canonical form (69) (or, equivalently (57)) is recovered immediately. It is for this reason that the form (57), which is valid whether or not the λ_i are distinct, was called the "Generalized Lur'e Canonical Form."

In a subsequent paper [11], an explicit, analytic, non-singular, nonlinear transformation

$$\sigma = g(\phi) = g(TL^*x) , \tag{127}$$

will be defined which transforms the Generalized Lur'e Form (57), for constant ψ_o, into the simplest possible canonical form, namely

$$\dot{\sigma} = \psi_o e^n . \tag{128}$$

The use of (57) in the form (125), which is valid whether or not the λ_i are distinct, is the key to a very direct proof of the important result (128).

SUMMARY

A. Major Definitions and Identities

For the system $\dot{x} = Ax + a\psi_o$, in general:

$$(sI - A)^{-1} = \frac{\Gamma(s)}{\Delta(s)} ,$$

$$\Delta(s) = \det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_o = 0 ,$$

$$\Gamma(s) = \sum_{i=1}^{n} s^{i-1}S_i ,$$

$$S_i = \sum_{j=i}^{n} \alpha_j A^{j-i} , \qquad (i = 0, 1, \cdots, n) , \qquad S_o \equiv 0 ,$$

$$D = (a, Aa, \cdots, A^{n-1}a) , \qquad \det D \neq 0 ,$$

$$D^*b = e^n ,$$

$$L = (b, A^*b, \cdots, (A^*)^{n-1}b) ,$$

$$L^{-1} = (S_1a, S_2a, \cdots, S_na)^* ,$$

$$(L^{-1})^* = DT ,$$

$$D^{-1} = TL^* = (S_1^*b, S_2^*b, \cdots, S_n^*b)^* ,$$

$$T = (t^1, t^2, \cdots, t^n) , \qquad t^i = \sum_{j=i}^{n} \alpha_j e^{j-i+1} , \quad (i = 1, 2, \cdots, n),$$

$$T^{-1} = (\tau^1, \tau^2, \cdots, \tau^n) ; \qquad \tau^i = \sum_{j=0}^{i-1} \beta_j e^{n+j-i+2} , \quad (i = 1, 2, \cdots, n),$$

$$\beta_\nu = \sum_{j=0}^{\nu-1} \alpha_{j+n-\nu}\beta_j \ , \qquad (\nu = 1, 2, \cdots, n) \ , \qquad \beta_o = 1 \ ,$$

$$C = \Big(-\alpha_o e^n, e^1 - \alpha_1 e^n, \cdots, e^{j-1}$$
$$-\alpha_{j-1} e^n, \cdots, e^{n-1} - \alpha_{n-1} e^n\Big) \ ,$$

$$L^*A(L^*)^{-1} = C \ ,$$

$$D^{-1}AD = C^* \ .$$

For n roots λ_i of $\Delta(s) = 0$ distinct:

$$Z = (z^1, z^2, \cdots, z^n) \ , \qquad z^i = \sum_{k=1}^{n} (\lambda_i)^{k-1} e^k$$

$$W = (w^1, w^2, \cdots, w^n) \ , \qquad w^i = Tz^i/\Delta'(\lambda_i)$$

$$= \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \sum_{k=j}^{n} \alpha_k e^{k-j+1} \ ,$$

$$W = (Z^*)^{-1}$$

$$V = (v^1, v^2, \cdots, v^n) \ , \qquad v^i = \sum_{j=1}^{n} (\lambda_i)^{j-1} S_j^* b \ ,$$

$$U = (u^1, u^2, \cdots, u^n) \ , \qquad u^i = Dw^i = \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} S_j a$$

$$U = (V^*)^{-1} \ , \qquad D^*V = Z \ ,$$

$$Au^i = \lambda_i u^i \ , \qquad u^i \cdot b = 1/\Delta'(\lambda_i) \ , \qquad (i = 1, 2, \cdots, n) \ ,$$

$$A^{*}v^{i} = \lambda_i v^{i}\,, \qquad v^{i}\cdot a = 1\,, \qquad (i = 1, 2, \cdots, n)\,,$$

$$b \equiv VU^{*}b = \sum_{i=1}^{n} v^{i}(u^{i})^{*}b = \sum_{i=1}^{n}\left\{\frac{1}{\Delta'(\lambda_i)}\right\}v^{i}\,.$$

B. Coordinate Transformations in Vector-Matrix Form

	x	θ	ϕ	ξ $(\lambda_i \neq \lambda_j)$
x	$x = x$	$\theta = L^{*}x$	$\phi = TL^{*}x$	$\xi = V^{*}x$
θ	$x = DT\theta$	$\theta = \theta$	$\phi = T\theta$	$\xi = Z^{*}T\theta$
ϕ	$x = D\phi$	$\theta = T^{-1}\phi$	$\phi = \phi$	$\xi = Z^{*}\phi$
ξ $(\lambda_i \neq \lambda_j)$	$x = DW\xi$	$\theta = T^{-1}W\xi$	$\phi = W\xi$	$\xi = \xi$

C. Coordinate Transformations in Vector-Scalar Form

	x	θ	ϕ	ξ $(\lambda_i \neq \lambda_j)$
x	$x_i = x_i$	$\theta_i = (A^{*})^{i-1}b\cdot x$	$\phi_i = (S_i^{*}b)\cdot x$	$\xi_i = v^{i}\cdot x$
θ	$x = \sum_{i=1}^{n}\theta_i S_i a$	$\theta_i = \theta_i$	$\phi_i = \sum_{j=i}^{n}\alpha_j\theta_{j-i+1}$	$\xi_i = \sum_{j=1}^{n}\left\{\sum_{k=j}^{n}\alpha_k\lambda_i^{k}\right\}\theta_j$
ϕ	$x = \sum_{i=1}^{n}\cdot\phi_i A^{i-1}a$	$\theta_i = \sum_{j=0}^{i-1}\beta_j\phi_{j+n-i+1}$	$\phi_i = \phi_i$	$\xi_i = \sum_{j=1}^{n}\lambda_i^{j-1}\phi_j$
ξ $(\lambda_i \neq \lambda_j)$	$x = \sum_{i=1}^{n}\xi_i u^{i}$	$\theta_i = \sum_{j=1}^{n}\frac{(\lambda_j)^{i-1}}{\Delta'(\lambda_j)}\xi_j$	$\sum_{j=1}^{n}\left\{\sum_{\nu=i}^{n}\alpha_\nu\frac{(\lambda_j)^{\nu-i}}{\Delta'(\lambda_j)}\right\}\xi_j$	$\xi_i = \xi_i$

D. Canonical Forms in Vector-Matrix Notation

$$\dot{x} = Ax + a\psi_o \ ,$$

$$\dot{\theta} = C\theta + e^n\psi_o \ ,$$

$$\dot{\phi} = C^*\phi + e^1\psi_o \ ,$$

$$\dot{\xi} = \Lambda\xi + u^o\psi_o \ , \qquad (u^o = e^1 + e^2 + \cdots + e^n) \ .$$

E. Canonical Forms in Vector-Scalar Notation

$$\Delta(d/dt)\theta_1 = \psi_o \ , \qquad \theta_1 = \phi_n \ ;$$

$$\dot{\xi}_o(s) = s\xi_o(s) + \psi_o - \Delta(s)\phi_n \ , \qquad \xi_o(s) = \sum_{i=1}^{n} s^{i-1}\phi_i \ ;$$

$$\dot{\xi}_i = \lambda_i\xi_i + \psi_o \ , \qquad \xi_i = \xi_o(\lambda_i) \ ,$$

for λ_i all distinct, $(i = 1, 2, \cdots, n)$.

REFERENCES

1. Lur'e, A.I., Some Non-Linear Problems in the Theory of Automatic Control, London, Her Majesty's Stationary Office, 1957.

2. Letov, A.M., Stability in Nonlinear Control Systems, Princeton, Princeton University Press, 1961.

3. La Salle, J. and Lefschetz, S., Stability by Liapunov's Direct Method with Applications, New York, Academic Press, 1961.

4. Bass, R.W. and Mendelson, P., Aspects of General Control Theory, Air Force Office of Scientific Research Report No. AFOSR2754, August 1962.

5. Lewis, D.C. and Mendelson, P., "Contributions to the Theory of Optimal Control", Transactions of the American Mathematical Society, Vol. 110, No. 2, pp. 232-244, February 1964.

6. Wonham, W.M. and Johnson, C.D., "Optimal Bang-Bang Control with Quadratic Performance Index", Transactions of the ASME-Journal of Basic Engineering, pp. 107-115, March 1964.

7. Wonham, W.M. and Johnson, C.D., "On a Problem of Letov in Optimal Control", Preprints of Papers - 1964 JACC, pp. 317-325, June 1964.

8. Wonham, W.M. and Johnson, C.D., "A Note on the Transformation to Canonical (Phase-Variable) Form", IEEE Transactions on Automatic Control, pp. 312-313, July 1964.

9. Kalman, R.E., "When is a Linear Control System Optimal?", Transactions of the ASME Journal of Basic Engineering, pp. 51-60, March 1964.

10. Bass, R.W. and Gura, I., "High Order System Design via State Space Considerations", Preprints of Papers 1965 JACC (to appear).

11. Bass, R.W. and Gura, I., "A Nonlinear Canonical Form for Controllable Systems" (to appear).

12. Bass, R.W. and Gura, I., "Synthesis of Time-Optimal Feedback Control Systems" (to appear).

13. Bass, R.W. and Webber, R.F., "On Synthesis of Optimal Bang-Bang Feedback Control Systems with Quadratic Performance Index", Preprints of Papers 1965 JACC (to appear).

14. Zadeh, L.A. and Desoer, C.A., Linear System Theory, New York, McGraw-Hill Book Company, Inc., pp. 303 ff., 1963.

APPENDIX B

A NONLINEAR CANONICAL FORM FOR CONTROLLABLE BANG-BANG SYSTEMS

by

R. W. Bass and I. Gura
Space Systems Division,
Hughes Aircraft Company
under

Contract No. NAS8-11421

Space Systems Division

AEROSPACE GROUP

Hughes Aircraft Company · Culver City, California

APPENDIX B
A NONLINEAR CANONICAL FORM FOR CONTROLLABLE BANG-BANG SYSTEMS

SUMMARY

Consider the state-vector control system $\dot{x} = Ax + a\epsilon$, $\epsilon = \pm 1$, where the pair (A, a) satisfies the condition of controllability. It is known from general existence therorems [2], [3] that there exists near $x = 0$, a nonlinear non-singular coordinate transformation $\sigma = p(x, \epsilon)$ such that the given system is equivalent to the simplest possible system, $\dot{\sigma} = e^n \epsilon$, $e^n = (0, 0, \cdots, 0, 1)^*$, whose state-space trajectories are parallel straight lines. Here the function $p(x, \epsilon)$, and its inverse $h(\sigma, \epsilon)$, where $\sigma \equiv p[h(\sigma, \epsilon), \epsilon]$, are defined explicitly by closed-form expressions involving only rational functions and the elementary transcendental functions. Various problems of stabilization and optimization can be solved in the σ-coordinates and the answers then applied to the original system in x-coordinates. In many cases [8], [9] it is possible to define scalar functions $\tilde{\sigma}(x)$ and $\psi(\sigma)$ such that the desired control law is given in the form $\epsilon = \mathrm{sgn}\{\psi[p(x, \tilde{\epsilon})]\}$, $\tilde{\epsilon} = \mathrm{sgn}[\tilde{\sigma}(x)]$ which is readily mechanizable by means of the explicit representation for p(x, t)

INTRODUCTION

In a previous paper [1] several linear coordinate transformations were defined such that useful canonical forms of the system differential equations can be easily obtained.

Here a nonlinear coordinate transformation is defined which changes any controllable linear bang-bang system into the simplest possible system, namely one whose state-space phase portrait consists of parallel straight lines. Evolution of the system in time then corresponds to uniform rectilinear motion.

The theory of integrals and isochrones [2], [3] will be reviewed in a general setting. Then for controllable linear systems a complete set of integrals and isochrones will be given by means of contour integrals in the complex s-plane [Equation (40)]. Alternate expressions suitable for use in computer-algorithms will be derived using Lur'e coordinates [Equation (45)], generalized Lur'e coordinates [Equations (66) and (67)], and phase coordinates [Equation (84)]. Because of the usefulness of these integrals and isochrones in designing and simulating optimal control systems, the algebraic and analytic details of their construction will be presented in full. It is assumed that the reader is somewhat familiar with the results of [1].

NOTATIONAL CONVENTIONS

a. Matrices are upper case Roman letters.

b. Vectors are lower case unsubscripted or superscripted Roman letters.

c. Scalars are Greek letters and all subscripted lower case letters.

d. Exceptions to these rules are as follows:

   1) i, j, k, l, ν, m, n are used as summation indices or scalars.

   2) θ, ϕ, ξ, σ (unsubscripted) are vectors.

   3) s is a complex scalar.

   4) $\Delta(s)$ is a scalar polynomial in s; $\Gamma(s)$ is a matrix polynomial in s.

   5) t is a scalar denoting time.

e. Asterisks used as superscripts (*) denote matrix transposition.

f. The i^{th} column of the identity matrix is represented by e^i.

g. The symbol $\triangleq$ denotes equality by definition; the symbol $\equiv$ denotes identity.

DEFINITIONS AND INTERPRETATION

A <u>first integral</u> of the n^{th} order system

$$\dot{x} = f(x), \quad x(0) = x^o, \tag{1}$$

is a scalar function $\sigma_*(x)$ such that

$$\sigma_*[x(t)] \equiv \sigma_*(x^o) \tag{2}$$

is satisfied along any solution of (1). Alternatively, $\sigma_*(x)$ can be defined by the condition

$$f(x) \cdot \text{grad}\ \sigma_*(x) \equiv 0 . \tag{3}$$

The equivalence† between (2) and (3) follows directly from the identity

$$\frac{d\sigma_*[x(t)]}{dt} \equiv \dot{x}(t) \cdot \text{grad}\ \sigma_*[x(t)] \equiv f(x)\cdot \text{grad}\ \sigma_*(x)\Big|_{x\,=\,x(t)} . \tag{4}$$

Geometrically, (2) defines an <u>integral surface</u> such that any state-space trajectory initiating on it, must remain on it for all t. Henceforth, the term "integral" will be used interchangeably for the function $\sigma_*(x)$ and the surface $\sigma_*(x) = \text{constant}$. The meaning should be clear from the context.

An <u>isochrone</u> is a surface defined by setting the scalar function $\sigma_o(x) = \text{constant}$ where $\sigma_o(x)$ satisfies

$$\sigma_o[x(t)] \equiv \sigma_o(x^o) + t \tag{5}$$

along any solution of (1). Note that, as in (2) - (3), the condition (5) is equivalent to

$$f(x) \cdot \text{grad}\ \sigma_o(x) \equiv 1 . \tag{6}$$

For a geometric interpretation, assume that two trajectories of (1) start on the same isochrone. Let the initial condition be x^o for one, and $\tilde{x}^o$ for the other. Then

$$\sigma_o(x^o) = \sigma_o(\tilde{x}^o) . \tag{7}$$

At some time t assume that the first trajectory crosses another isochrone defined by $\sigma_o[x(t)]$. Let $\tilde{t}$ represent the time at which the second trajectory crosses this isochrone. Then by definition

---

†Clearly, it is only necessary that (3) hold identically on the single surface $\sigma_*(x) = \sigma_*(x^o)$; however, if (3) holds in a neighborhood of x^o, then there exists a family of integral surfaces $\sigma_*(x) = \text{constant}$ in that neighborhood.

$$\sigma_o[x(t)] = \sigma_o[\tilde{x}(t)] \, . \tag{8}$$

The characteristic property of an isochrone is such that (8) must imply

$$t = \tilde{t} \, . \tag{9}$$

Thus the time for points on various trajectories to move between fixed isochrones is constant; hence the term "isochrone." In subsequent work this term will refer to either the function $\sigma_o(x)$ or the surface $\sigma_o(x) = \text{constant}$.

A <u>regular point</u> $\hat{x}$ is one such that $f(\hat{x}) \neq 0$. A <u>singular point</u> $\hat{x}$, which is such that $f(\hat{x}) = 0$, provides an <u>equilibrium solution</u> $x(t) \equiv \hat{x}$ of (1).

GENERAL THEORY OF INTEGRALS AND ISOCHRONES

<u>Theorem 1</u>. If $\sigma_1, \sigma_2, \cdots, \sigma_{n-1}$ are integrals for (1), then so is

$$\sigma_*(x) = \zeta(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) \, , \tag{10}$$

where ζ is an arbitrary function of its n-1 arguments.

<u>Proof</u>. By composite differentiation, (10) yields

$$f(x) \cdot \text{grad}\, \sigma_*(x) = \sum_{j=1}^{n-1} \frac{\partial \zeta}{\partial \sigma_j} [f(x) \cdot \text{grad}\, \sigma_j] = 0 \, . \tag{11}$$

<u>Theorem 2</u>. Every integral $\sigma_*(x)$ can be expressed in the form (10) in a neighborhood of a regular point $\hat{x}$ if $\sigma_1, \sigma_2, \cdots, \sigma_{n-1}$ are functionally independent at that point [i. e., the vectors grad σ_i, $(i = 1, 2, \cdots, n-1)$, evaluated at $\hat{x}$, are linearly independent].

<u>Proof</u>. If the n-vectors grad σ_i, $(i = 1, 2, \cdots, n-1)$, are linearly independent at $\hat{x}$, then the matrix

$$
\begin{bmatrix}
\frac{\partial \sigma_1}{\partial x_1} & \frac{\partial \sigma_2}{\partial x_1} & \cdots & \frac{\partial \sigma_{n-1}}{\partial x_1} \\
\frac{\partial \sigma_1}{\partial x_2} & \frac{\partial \sigma_2}{\partial x_2} & \cdots & \frac{\partial \sigma_{n-1}}{\partial x_2} \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\frac{\partial \sigma_1}{\partial x_{n-1}} & \frac{\partial \sigma_2}{\partial x_{n-1}} & \cdots & \frac{\partial \sigma_{n-1}}{\partial x_{n-1}}
\end{bmatrix}_{(\hat{x})}, \qquad (12)
$$

must be non-singular. [Note that since the x's can be arranged arbitrarily, x_n can be chosen with no loss of generality as that variable for which the vector $\left(\frac{\partial \sigma_1}{\partial x_n}, \frac{\partial \sigma_2}{\partial x_n}, \cdots, \frac{\partial \sigma_{n-1}}{\partial x_n}\right)$ is linearly dependent on the rows of (12)]. Then by the Implicit Function Theorem [see Appendix 1] the transformation

$$\sigma_i = p_i(x_1, x_2, \cdots, x_{n-1}, x_n), \qquad (i = 1, 2, \cdots, n-1) \qquad (13)$$

has a unique inverse

$$x_k = h_k(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}, x_n), \qquad (k = 1, 2, \cdots, n-1) \qquad (14)$$

in a neighborhood of $\hat{x}$.

Or, in that neighborhood,

$$\sigma_i \equiv p_i[h_1(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}, x_n), \cdots, h_{n-1}(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}, x_n), x_n]. \quad (15)$$

Since (15) is an identity, σ_i must be independent of x_n; hence

$$\frac{\partial \sigma_i}{\partial x_n} \equiv 0 \equiv \left[\left(\sum_{k=1}^{n-1} \frac{\partial p_i}{\partial x_k} \frac{\partial h_k}{\partial x_n} \right) + \frac{\partial p_i}{\partial x_n} \right], \quad (i=1, \cdots, n) . \tag{16}$$

Now consider an arbitrary integral

$$\sigma_* = \zeta(x_1, x_2, \cdots, x_{n-1}, x_n). \tag{17}$$

Applying (14),

$$\sigma_* = p_*[h_1(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}, x_n), \cdots, h_{n-1}(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}, x_n), x_n], \tag{18}$$

and so

$$\frac{d\sigma_*}{dx_n} = \left[\left(\sum_{k=1}^{n-1} \frac{\partial \zeta}{\partial x_k} \frac{\partial h_k}{\partial x_n} \right) + \frac{\partial \zeta}{\partial x_n} \right] . \tag{19}$$

Before proceeding, note that the definition of an integral requires that

$$f(x) \cdot \text{grad } \sigma_* \equiv 0 , \quad f(x) \cdot \text{grad } \sigma_i \equiv 0 , \quad (i=1, 2, \cdots, n-1) , \tag{20a}$$

or, in vector-matrix form,

$$[\text{grad } \sigma_1, \text{ grad } \sigma_2, \cdots, \text{ grad } \sigma_{n-1}, \text{grad } \sigma_*]^* f(x) \equiv 0 . \tag{20b}$$

Since $f(x) \neq 0$, this can only be valid if, at $\hat{x}$,

$$\text{grad } \sigma_* = \sum_{i=1}^{n-1} \gamma_i \text{ grad } \sigma_i , \tag{21}$$

where the γ_i are constants not all zero. In scalar notation, (21) is

$$\frac{\partial \zeta}{\partial x_k} = \sum_{i=1}^{n-1} \gamma_i \frac{\partial p_i}{\partial x_k}, \qquad (k=1, 2, \cdots, n), \tag{22}$$

With this, (19) becomes

$$\frac{d\sigma_*}{dx_n} = \left(\sum_{k=1}^{n-1} \sum_{i=1}^{n-1} \gamma_i \frac{\partial p_i}{\partial x_k} \frac{\partial h_k}{\partial x_n} \right) + \left(\sum_{i=1}^{n-1} \gamma_i \frac{\partial p_i}{\partial x_n} \right), \tag{23}$$

or, rearranging the summations,

$$\frac{d\sigma_*}{dx_n} = \sum_{k=1}^{n-1} \gamma_i \left[\left(\sum_{i=1}^{n-1} \frac{\partial p_i}{\partial x_k} \frac{\partial h_k}{\partial x_n} \right) + \frac{\partial p_i}{\partial x_n} \right]. \tag{24}$$

By (16), then,

$$\frac{d\sigma_*}{dx_n} \equiv 0, \tag{25}$$

which indicates that σ_* is <u>not</u> a function of x_n. Thus the construction (18) has defined a function ζ such that

$$\sigma_* = \zeta(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}), \tag{26}$$

and the proof of the theorem is complete.†

<u>Theorem 3</u>. Let σ_* be an integral, and σ_n an isochrone. Then the function

$$\sigma_o = \sigma_* + \sigma_n \tag{27}$$

is an isochrone.

†A theorem similar to Theorem 2 is given in [4, p. 115]. However, that theorem refers to n independent time-varying integrals; in [4], if σ_n is an isochrone, $(\sigma_n - t)$ is called an integral. Theorem 2 is not a direct corollary of [4, p. 115].

Proof. Since

$$f(x) \cdot \operatorname{grad} \sigma_o = f(x) \cdot \operatorname{grad} \sigma_* + f(x) \cdot \operatorname{grad} \sigma_n = 0 + 1 = 1 , \tag{28}$$

(27) must be an isochrone.

Theorem 4. Let σ_n be an isochrone, and let $\sigma_1, \sigma_2, \cdots, \sigma_{n-1}$ be first integrals, functionally independent at a regular point $\hat{x}$. Then every isochrone σ_o can be expressed in the form

$$\sigma_o = \zeta(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) + \sigma_n \tag{29}$$

in a neighborhood of a point $\hat{x}$ for an appropriate function ζ.

Proof. By hypothesis,

$$f(x) \cdot \operatorname{grad} \sigma_o \equiv 1, \; f(x) \cdot \operatorname{grad} \sigma_n \equiv 1 \tag{30}$$

whence

$$f(x) \cdot \operatorname{grad} (\sigma_o - \sigma_n) \equiv 0 , \tag{31}$$

and $\sigma_o - \sigma_n \triangleq \sigma_*$ must be an integral. In a neighborhood of $\hat{x}$ Theorem 2 applies, and σ_* must have the form (10). Thus every isochrone must have the form (29) in that neighborhood.

RELATIONSHIP TO CONTROL THEORY

The value of the concepts introduced in the preceding pages to the theory of automatic controls is embodied in the following theorem:

Theorem 5. If $\sigma_1(x;\epsilon)$, $\sigma_2(x;\epsilon), \cdots, \sigma_{n-1}(x;\epsilon)$ are first integrals and $\epsilon\sigma_n(x;\epsilon)$ is an isochrone for the system

$$\dot{x} = Ax + a\epsilon\,, \quad (\epsilon^2 = 1)\,, \tag{32}$$

and if the elements of the vector $\sigma = [\sigma_1(x;\epsilon), \sigma_2(x;\epsilon), \cdots, \epsilon\sigma_n(x;\epsilon)]$ are functionally independent at $\hat{x}$, then in a neighborhood of $\hat{x}$ there exists a unique transformation $\sigma = p(x;\epsilon)$, and inverse $x = h(\sigma;\epsilon)$ between the system (32) and the system

$$\dot{\sigma} = \epsilon e^n\,, \tag{33}$$

where p and h are n-vector functions and ϵ is a parameter only taking on the values +1 or -1. (The reason for making ϵ a factor in the definition of the isochrone will become evident later.)

Proof. By the definitions of integrals and isochrones, $\dot{\sigma}_i = 0$, $(i = 1, \cdots, n-1)$ and $\epsilon\dot{\sigma}_n = 1$, or $\dot{\sigma}_n = \epsilon$, whence (32) implies (33). However, by the Implicit Function Theorem [see Appendix 1] , the implicit equation $\sigma - p(x, \epsilon) = 0$ has a unique solution $x = h(\sigma;\epsilon)$ near $\hat{x}$, because the Jacobian matrix $\partial(\sigma - p)_i/\partial x_j = -\partial p_i/\partial x_j$ is non-singular at $\hat{x}$ by hypothesis.

Geometrically, the nonlinear change of coordinates described above rectifies the state-space flow of (32) into the most elementary possible dynamical system, namely uniform rectilinear motion along parallel straight lines. Solution of the system (33) is, of course, trivial.

APPLICATION TO HAMILTON-JACOBI EQUATION

If the transformation between (32) and (33) (i.e., between x and σ coordinates) can be found explicitly, solution of the Hamilton-Jacobi partial differential equation encountered in optimal control theory is facilitated.

Consider the problem of choosing the control ϵ in (33) such that the cost functional (or performance index)

$$\Phi(x^o) = \int_o^{t_1} \Psi(x)\, dt \tag{34}$$

is minimized. Here $x(t) = x^1$ is a given stopping condition so that the terminal time $t_1 = t_1(x^o)$. The optimal control $\epsilon = \epsilon(x)$ then is given by the solution of the Hamilton-Jacobi Equation [6], [7]

$$\min_{\epsilon} \left\{ x \cdot A^* \operatorname{grad} \Phi(x) + \epsilon\, a \cdot \operatorname{grad} \Phi(x) + \Psi(x) \right\} = 0 \ . \tag{35}$$

When $a \cdot \operatorname{grad} \Phi \neq 0$, this expression is minimized by the choice

$$\epsilon = -\operatorname{sgn}[a \cdot \operatorname{grad} \Phi(x)] \ , \tag{36a}$$

and so (35) becomes

$$x \cdot A^* \operatorname{grad} \Phi(x) - \left| a \cdot \operatorname{grad} \Phi(x) \right| + \Psi(x) = 0 \ . \tag{36b}$$

Now, when Theorem 5 applies, Equation (36) may be transformed from x-coordinates to σ-coordinates by setting

$$\Phi(x) = \Phi\, [h(\sigma;\epsilon)] \triangleq \hat{\Phi}(\sigma;\epsilon) \ , \tag{37a}$$

$$\Psi(x) = \Psi[h(\sigma;\epsilon)] \triangleq \hat{\Psi}(\sigma;\epsilon) \ . \tag{37b}$$

Correspondingly, the pair (A, a) becomes $(0, e^n)$, and so the Hamilton-Jacobi Equation (36) becomes

$$\epsilon \frac{\partial \hat{\Phi}(\sigma;\epsilon)}{\partial \sigma_n} = -\hat{\Psi}(\sigma;\epsilon) \ , \qquad \epsilon = -\operatorname{sgn}\left[\frac{\partial \hat{\Phi}}{\partial \sigma_n}\right] \ . \tag{38}$$

For constant ϵ, the general solution of (38) is

$$\hat{\Phi}(\sigma;\epsilon) = \hat{\Phi}_o(\sigma_1, \sigma_2, \cdots, \sigma_{n-1};\epsilon) - \epsilon \int_o^{\sigma_n} \hat{\Psi}(\sigma, \epsilon)\, d\sigma_n \ , \tag{39}$$

where $\hat{\Phi}_o(\sigma_1, \sigma_2, \cdots, \sigma_{n-1};\epsilon)$ is an arbitrary function. In fact, $\partial\hat{\Phi}/\partial\sigma_n \equiv 0$, whence (39) obviously satisfies (38). Thus (39) is a particular solution of (39). On the other hand, the difference between any two particular solutions of (38) must be a solution of $\partial\hat{\Phi}/\partial\sigma_n$, which is fully accounted for by the arbitrariness of $\hat{\Phi}_o$ in (39).

Thus if the transformation laws $x = h(\sigma;\epsilon)$ and $\sigma = p(x;\epsilon)$ are known, an important class of optimal control problems can be reduced explicitly to the problem of properly piecing together functions of the type (39).

EXPLICIT CLOSED-FORM TRANSFORMATION FROM x TO σ

Theorem 6. The system (32) is equivalent to (33) under the transformation defined by

$$\xi_o(s) = v(s)\cdot x \tag{40a}$$

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} \frac{\epsilon}{s} \log\left[1 + \epsilon s \xi_o(s)\right] ds, \quad (j = 1, 2, \cdots, n) \tag{40b}$$

where $||x||$ is sufficiently small so that $|s\xi_o(s)| \le 1$, and the path of integration is a circle enclosing all the roots of $\Delta(s) = \Delta(\lambda_i) = 0$, $(\rho > \max|\lambda_i|)$. (Recall that the quantities v(s) and $\Delta(s)$ are defined in in [1].)

Proof. Differentiate (40b) with respect to time and apply [1, (125)]† to obtain

$$\dot{\sigma}_j = \frac{\epsilon}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} \frac{\epsilon}{1+\epsilon s\xi_o(s)} \left[s\xi_o(s) + \epsilon - \Delta(s)\theta_1\right] ds \quad , \qquad (41a)$$

or

$$\dot{\sigma}_j = \frac{\epsilon}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} \frac{\left[\epsilon s\xi_o(s) + \epsilon^2\right] ds}{1+\epsilon s\xi_o(s)} - \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}\theta_1 \, ds}{1+\epsilon s\xi_o(s)} \quad . \qquad (41b)$$

Now, since $\epsilon = \pm 1$, the first term of the right side of (41b) becomes

$$\frac{\epsilon}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} = \epsilon\delta_{jn} \quad , \qquad (j = 1, 2, \cdots, n) \quad . \qquad (42)$$

The derivation of this result is given in Appendix 2. The remaining term on the right hand side of (41b) can be expressed as

$$-\frac{\theta_1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{1+\epsilon s\xi_o(s)} = \frac{\theta_1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} s^{j-1} \left\{ \sum_{k=0}^{\infty} (-1)^k \left[\epsilon s\xi_o(s)\right]^k \right\} ds \equiv 0, \qquad (43)$$

if $\left|\epsilon\rho\xi_o(\rho)\right| = \left|\rho\xi_o(\rho)\right| < 1$. (In fact, since the integrand in (43) is analytic in s, Cauchy's Theorem implies that the integral in (43) is identically zero.) Combining (42) and (43) with (41b) yields

$$\dot{\sigma}_j = \delta_{jn} \quad , \qquad (j = 1, 2, \cdots, n) \quad , \qquad (44a)$$

or, in vector notation, the desired system equation

$$\dot{\sigma} = \epsilon e^n \qquad (44b)$$

must be valid. Note that the condition $\left|\rho\xi_o(\rho)\right| < 1$ can be obtained as a constraint on $\|x\|$ by applying (40a) to obtain $\|x\| < \frac{1}{\rho\bar{v}(\rho)}$ where $\bar{v}$ is the upper bound of $\|v(s)\|$ on $|s| = \rho$.

†[1, (125)] refers to Reference [1], Equation (125).

Corollary 6.1. For distinct λ_i, $(i = 1, 2, \cdots, n)$, it is clear that integration of (40b) by the Calculus of Residues yields the closed-form expressions

$$\xi_i = v^i \cdot x \,, \quad v^i = v(\lambda_i) \,, \quad (i = 1, 2, \cdots, n) \,, \tag{45a}$$

$$\sigma_j = \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \left\{ \frac{\epsilon}{\lambda_i} \log \left[1 + \epsilon \lambda_i \xi_i \right] \right\} , \quad (j = 1, 2, \cdots, n) \,, \tag{45b}$$

where ξ_i are simply the components of the state vector in Lur' e canonical form [1].

In cases for which the system eigenvalues are non-distinct, the explicit evaluation of σ is not as simple as in (45). For convenience, define

$$\eta(s, \xi, \epsilon) = \frac{\epsilon s^{j-1}}{s} \log \left[1 + \epsilon s \xi_o(s) \right] . \tag{46}$$

Then (42) becomes

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{\eta(s, \xi_o, \epsilon)}{\Delta(s)} \, ds \,. \tag{47}$$

Assume that $\Delta(s) = 0$ has $\ell \leq n$ distinct roots λ_i such that λ_i is a root of multiplicity j_i, that is,

$$\Delta(s) = (s - \lambda_1)^{j_1} (s - \lambda_2)^{j_2} \cdots \cdot (s - \lambda_\ell)^{j_\ell} \,, \tag{48}$$

where

$$j_1 + j_2 + \cdots j_m = n \,. \tag{49}$$

Then by a partial fraction expansion in (47), [5],

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \left\{ \sum_{i=1}^{\ell} \sum_{\nu=1}^{j_i} \frac{k_{i\nu}}{(s - \lambda_i)^{j_i - \nu + 1}} \right\} ds \quad , \tag{50}$$

where ℓ is the number of distinct roots of $\Delta(s) = 0$ and

$$k_{i\nu} \triangleq \frac{1}{(\nu - 1)!} \left\{ \frac{d^{\nu-1}}{ds^{\nu-1}} \left[\frac{(s - \lambda_i)^{j_i} \eta(s, \xi_o, \epsilon)}{\Delta(s)} \right] \right\}_{s = \lambda_i} \quad . \tag{51}$$

The theory of complex integration then yields the following result.

Corollary 6.2. For non-distinct λ_i, the transformation (40) has the closed-form expression

$$\sigma_j = \sum_{i=1}^{\ell} [k_{i\nu}]_{\nu = j_i} \quad , \quad (j = 1, 2, \cdots, n) \; . \tag{52}$$

EXPLICIT CLOSED-FORM TRANSFORMATION FROM σ TO x

Theorem 7. Assume that the system (32) is controllable. The transformation $\sigma = p(x, \epsilon)$ has a unique inverse $x = h(\sigma, \epsilon)$ given by

$$x = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{\epsilon}{s} \left\{ \exp\left[\epsilon \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} s^{\ell+1} \sigma \right] - 1 \right\} u(s) \, ds \quad , \tag{53}$$

where u(s) is defined in [1].

Corollary 7.1. When the λ_i are distinct, the inverse of (45) is given by

$$x = \sum_{i=1}^{n} \frac{\epsilon}{\lambda_i} \left\{ \exp\left[\epsilon \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} \lambda_i^{\ell+1} \sigma_\nu \right] - 1 \right\} u^i \tag{54}$$

Proof of Theorem 7 and Corollary 7.1. The simplest proof of Theorem 7 seems to be that in which Corollary 7.1 is proved first, independently, and then used as a lemma in the establishment of the theorem. In other words, (54) will be proved and then generalized to (53); subsequently, (54) can be recovered as a special case of (53).

Consider (45) and define a vector q such that each component is given by

$$q_i = \frac{\epsilon}{\Delta'(\lambda_i)} \frac{1}{\lambda_i} \log \left| 1 + \epsilon\lambda_i \xi_i \right| \ , \qquad (i = 1, 2, \cdots, n) \ . \tag{55}$$

Then (45) can be expressed in vector-matrix form as

$$\sigma = Zq \ , \tag{56}$$

where Z is the Vandermonde Matrix.

As shown in [1], the inverse of Z is given by the transpose of a matrix $W = (w^1, w^2, \cdots, w^n)$ such that

$$w^i = \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \sum_{k=j}^{n} \alpha_k e^{k-j+1} \ , \qquad (i = 1, 2, \cdots, n) \ , \tag{57}$$

where the α_k are the coefficients of the characteristic polynomial $\Delta(s)$ of the system. Thus (56) yields

$$q = W^* \sigma \ , \tag{58a}$$

or

$$q_i = w^i \cdot \sigma \ , \qquad (i = 1, 2, \cdots, n) \ . \tag{58b}$$

Then, combining this with (5) gives

$$\xi_i = \frac{\epsilon}{\lambda_i} \left\{ \exp\left[\epsilon \lambda_i \Delta'(\lambda_i) \ w^i \cdot \sigma \right] - 1 \right\} \ . \tag{59}$$

Now, applying (57),

$$\epsilon\lambda_i \Delta'(\lambda_i)\, w^i \cdot \sigma = \epsilon \sum_{j=1}^{n} \sum_{k=j}^{n} \alpha_k \lambda_i^j e^{k-j+1} \cdot \sigma = \epsilon \sum_{j=1}^{n} \sum_{k=j}^{n} \alpha_k \lambda_i^j \sigma_{k-j+1} \cdot \sigma \; . \tag{60}$$

To transform this last formula to a more convenient form, replace k by a new index $\nu = k - j + 1$ and obtain

$$\epsilon\lambda_i \Delta'(\lambda_i)\, w^i \cdot \sigma = \sum_{j=1}^{n} \sum_{\nu=1}^{n-j+1} \alpha_{\nu+j-1} \lambda_i^j \sigma_\nu \; . \tag{61}$$

Interchanging the order of summation in (61) and letting $\ell = j - 1$,

$$\epsilon\lambda_i \Delta'(\lambda_i)\, w^i \cdot \sigma = \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} \lambda_i^{\ell+1} \sigma_\nu \; . \tag{62}$$

Then, applying this result and [1, (108)] to (59), it is clear that the desired transformation formula is

$$x = \sum_{i=1}^{n} \xi_i u^i = \sum_{i=1}^{n} \frac{\epsilon}{\lambda_i} \left\{ \exp\left[\epsilon \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} \lambda_i^{\ell+1} \sigma_\nu \right] - 1 \right\} u^i \; . \tag{63}$$

Now define a transformation $x = h(\sigma, \epsilon)$ by (53). Using the Calculus of Residues, it is clear that (63) is equivalent to (53) when the λ_i are distinct. Also, for distinct λ_i, (45) and (40) are equivalent. Hence it is certain that (53) is the inverse of (40), at least when the λ_i are distinct. It will now be shown that this proposition is valid for all systems, even when the λ_i are non-distinct. To verify this, consider (40b) in the

$$\sigma = p(x; \epsilon, A, a,) \quad , \tag{64}$$

and define A to be <u>simple</u> when the roots of its characteristic polynomial $\Delta(s)$ are distinct. It is well known that if A is not simple there are simple matrices A_o such that $\|A - A_o\|$ is arbitrarily small.

It has been shown that there exists a function $h(\sigma;\epsilon, A, a)$, namely (53), such that

$$h(p(x;\epsilon, A, a);\epsilon, A_\nu)) \triangleq \hat{h}(x;\epsilon, A, a) \equiv x \tag{65}$$

is valid whenever A is simple. Now take A non-simple. Let $\{A_\nu\}$ be a sequence such that A_ν is simple for each $\nu = 1, 2, 3, \cdots$ and such that $A_\nu \rightarrow A$ as $\nu \rightarrow \infty$. Now the integrand in (40b) is a continuous function of x, A, a, and ϵ since v(s) is a polynomial in A, a, and $1/\Delta(s)$ [1, (119b)]. Recall also that $1/\Delta(s)$ is an infinite series in powers of s^{-1}, which converges for $|s| > \max(\lambda_i)$, whose coefficients are rational functions of A. Thus $p(x;\epsilon, A, a)$ is a continuous function of all its arguments. Clearly, an analogous result can be obtained for $h(\sigma;\epsilon, A, a)$. Thus $\hat{h}(x,\epsilon,A,a)$ is continuous in all arguments and so $\hat{h}(x;\epsilon, A_\nu, a) \rightarrow \hat{h}(x;\epsilon, A, a)$ as $\nu \rightarrow \infty$. But since $\hat{h}(x;\epsilon, A_\nu, a) \equiv x$, it follows upon taken the limit that $\hat{h}(x;\epsilon, A, a) = x$. This completes the proof of the theorem.

EXPANSION OF σ IN SERIES OF RECURSIVELY COMPUTABLE MULTINOMIALS

Theorem 8. The functions σ_i defined in (40) may be expressed as

$$\sigma_i = \sum_{\ell=0}^{\infty} \beta_\ell \omega_{\ell+n-i+1} , \qquad (i = 1, 2, \cdots, n) \tag{66}$$

where the sequence of numbers $\{\beta_\ell\}$ is recursively computable from the definitions

$$\beta_o = 1, \qquad \beta_\ell = - \sum_{j=0}^{\ell-1} \alpha_{j+n-\ell} \beta_j , \qquad (\ell = 1, 2, \cdots, n) \tag{67a}$$

$$\beta_{\ell+n} = - \sum_{j=\ell}^{\ell+n-1} \alpha_{j-\ell} \beta_j , \quad (\ell = 1, 2, \cdots) \tag{67b}$$

and where the functions $\omega_\nu = \omega_\nu(x)$ are multinomials of degree ν in ϕ_1, $\phi_2, \cdots, \phi_n$, also recursively computable by

$$\omega_1 = \phi_1 \tag{68a}$$

$$\omega_\nu = \phi_\nu - \frac{\epsilon}{\nu} \sum_{m=1}^{\nu-1} m\, \omega_m \phi_{\nu-m} , \qquad (\nu = 2, \cdots, n) \tag{68b}$$

$$\omega_{\nu+n} = - \frac{\epsilon}{\nu+n} \sum_{i=1}^{n} (\nu+n-i) \phi_i \omega_{\nu+n-i} , \qquad \begin{matrix} (\nu = 1, 2, \cdots) , \\ (i = 1, 2, \cdots, n) \end{matrix} \tag{68c}$$

and the ϕ_i 's are linear functions of x defined in [1].

Proof. By a Taylor expansion

$$\epsilon \log (1+\epsilon s\xi_o(s)) = \sum_{j=1}^{\infty} \epsilon_o \frac{(-1)^{j+1}}{j} [\epsilon_o s\xi_o(s)]^j \quad , \tag{69}$$

for $|\epsilon s\xi_o(s)| < 1$. Now, since $\xi_o(s)$ is a polynomial in s, (40a), [1, (119b)], the right side of (69) is an infinite series in s and so

$$\epsilon \log [1+\epsilon s\xi_o(s)] = \sum_{j=1}^{\infty} \omega_j s^j \quad , \tag{70}$$

where the coefficients ω_j, $(j = 1, 2, \cdots)$, are to be determined. To accomplish this end, differentiate (70) with respect to s, obtaining

$$\frac{\epsilon^2 d[s\xi_o(s)]/ds}{1+\epsilon s\xi_o(s)} = \sum_{j=1}^{\infty} j\omega_j s^{j-1} \quad . \tag{71}$$

However, from [1, (123)]

$$\xi_o(s) = \sum_{i=1}^{n} s^{i-1}\phi_i \quad , \tag{72}$$

and so (71) becomes

$$\sum_{i=1}^{n}{}' i s^{i-1}\phi_i = \left[1+\epsilon\sum_{i=1}^{n} s^i\phi_i\right] \sum_{j=1}^{\infty} j\omega_j s^{j-1} = \tag{73a}$$

$$= \sum_{j=1}^{\infty} j\omega_j s^{j-1} + \epsilon \sum_{i=1}^{n}\sum_{j=1}^{\infty} j\omega_j\phi_i s^{i+j-1} \quad . \tag{73b}$$

Let k = i+j in the second sum on the right-hand side of (73b) and replace the index j by k-i to obtain

$$\sum_{i=1}^{n} is^{i-1}\phi_1 = \sum_{j=1}^{\infty} j\omega_j s^{j-1} + \epsilon \sum_{i=1}^{n} \sum_{k=i+1}^{\infty} (k-i)\phi_i\omega_{k-i}s^{k-1} \quad . \tag{74}$$

Then, interchanging the order of summation for the terms in question

$$\sum_{i=1}^{n} is^{i-1}\phi_1 = \sum_{j=1}^{\infty} j\omega_j s^{j-1} + \epsilon \sum_{k=2}^{\infty} \sum_{i=1}^{\substack{\min \\ (k-1,\, n)}} (k-i)\phi_i\omega_{k-1}s^{k-1} \quad . \tag{75}$$

Now, equating like coefficients of s in (75),

$$\omega_1 = \phi_1 \quad , \tag{76a}$$

$$\omega = \phi - \frac{\epsilon}{\nu} \sum_{i=1}^{\substack{\min \\ (\nu-1,\, n)}} (\nu-i)\phi_i\omega_{\nu-i} \quad ; \qquad (\nu = 1, 2, \cdots) \quad . \tag{76b}$$

For $\nu = 2, 3, \cdots, n$, let $m = \nu - i$ and replace i in (76b). Then

$$\omega_\nu = \phi_\nu - \frac{\epsilon}{\nu} \sum_{m=1}^{\nu=1} m\omega_m\phi_{\nu-m} \quad , \qquad (\nu = 2, \cdots, n) \quad . \tag{77a}$$

For $\nu = n+1, \cdots$, replace ν in (76b) by $\nu+n$ and obtain

$$\omega_{\nu+n} = - \frac{\epsilon}{\nu+n} \sum_{i=1}^{n} (\nu+n-i)\phi_i\omega_{\nu+n-i} \quad , \qquad (\nu = 1, 2, \cdots) \quad . \tag{77b}$$

Thus the ω's can be generated recursively as functions of the ϕ's as claimed.

Note that, using (70), (40b) can be expressed now by

$$\sigma_i = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{i-1}}{\Delta(s)} \frac{1}{s} \sum_{j=1}^{\infty} \omega_j s^j \, ds . \tag{78}$$

As shown in [1, (29)],

$$\frac{1}{\Delta(s)} = \sum_{\ell=0}^{\infty} \beta_\ell \, s^{-(n+\ell)} \; ; \qquad (|s| > \rho) \tag{79}$$

where the β's obey (67a, b, c). Then (78) becomes

$$\sigma_i = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \sum_{\ell=0}^{\infty} \sum_{j=1}^{\infty} \beta_\ell \omega_j s^{-n-\ell+j+i-2} ds, \qquad (i = 1, 2, \cdots, n) \quad . \tag{80a}$$

By residues, this becomes $[-n-\ell+j+i-1 = 0 \;\; \text{when} \;\; j = \ell+n+1-i]$

$$\sigma_i = \sum_{\ell=0}^{\infty} \beta_\ell \omega_{\ell+n+1-i} \;, \qquad (i = 1, 2, \cdots, n) \;, \tag{80b}$$

the desired result.

Corollary 8.1. The nth order scalar differential equation

$$\frac{d^n \theta_1}{dt^n} \equiv \theta_1^{[n]} = \epsilon \tag{81}$$

has for a complete system of integrals and an isochrone the multinomials

$$\sigma_i = \sigma_i\left(\theta_1^{[i-1]}, \theta_1^{[i]}, \cdots, \theta_1^{[n-1]}; \epsilon\right), \tag{82}$$

defined recursively by

$$\sigma_n = \theta_1^{[n-1]} \tag{83a}$$

$$\sigma_i = \theta_1^{[i-1]} - \frac{\epsilon}{n-i+1} \sum_{m=1}^{n-i} m\sigma_{n-m+1}\theta_1^{[i+m-1]} , \quad (i = 1, 2, \cdots, n-1) \tag{83b}$$

Proof. Since the characteristic equation for (81) is $\Delta(s) = s^n$, $\alpha_o = \alpha_1 = \cdots = \alpha_{n-1} = 0$. Then from (67a, b) it is clear that $\beta_\ell = 0$, $(\ell = 1, 2, \cdots)$, and so (66) becomes $\sigma_i = \omega_{n-i+1}$. Also, by [1, (65)] $\phi_i = \theta_{n-i+1} = \theta^{[n-i]}$, $(i = 1, 2, \cdots, n)$. Thus (68a) yields (83a) and (68b) yields (83b), directly.

The integrals of $\theta_1^{[3]} = \epsilon$ and $\theta_1^{[4]} = \epsilon$ given in [3] can be generated systematically by use of (83).

CLOSED-FORM EXPRESSIONS FOR COEFFICIENTS OF POWER SERIES EXPANSION OF σ

<u>Theorem 9.</u> The functions σ_i defined in (40) may be expressed as

$$\sigma_j = (A^*)^{j-1}\, b\cdot x - \frac{1}{2}\epsilon\,(x\cdot Q_j x) + \cdots, \quad (j = 1, 2, \cdots, n) \tag{84a}$$

where (see [1])

$$Q_1 = (D^{-1})^* \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 & \beta_1 \\ 0 & 0 & \cdots & 1 & \beta_1 & \beta_2 \\ 0 & 0 & \cdots & \beta_1 & \beta_2 & \beta_3 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 1 & \beta_1 & \cdots & \beta_{n-3} & \beta_{n-2} & \beta_{n-1} \\ \beta_1 & \beta_2 & \cdots & \beta_{n-2} & \beta_{n-1} & \beta_n \end{bmatrix} D^{-1} \tag{84b}$$

$$D^{-1} = (S_1^* b,\ S_2^* b,\ \cdots,\ S_n^* b), \quad S_i = \sum_{j=i}^{n} \alpha_j A^{j-i}, \tag{84c}$$

$$D = (a,\ Aa,\ \cdots,\ A^{n-1}a), \quad D^* b = e^n. \tag{84d}$$

$$Q_{j+1} = A^* Q_j\,, \quad (j = 1, 2, \cdots, n-1)\ . \tag{84e}$$

<u>Proof.</u> It is well known that for $\left|\lambda_i \xi_i\right| < 1$,

$$\frac{\epsilon}{\lambda_i}\log(1 + \epsilon\,\lambda_i\,\xi_i) = \epsilon^2 \xi_i - \frac{1}{2}\epsilon\lambda_i\,\xi_i^2 + \cdots. \tag{85}$$

For simplicity, assume (temporarily) that the λ_i are distinct. Then (45) becomes

$$\sigma_j = \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \xi_i - \frac{1}{2}\epsilon \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \lambda_i \xi_i^2 + \cdots. \tag{86}$$

By [1, (98)] and [1, (49a)]

$$\sigma_j = (A^*)^{j-1} b \cdot x - \frac{1}{2}\epsilon \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \lambda_i \xi_i^2 + \cdots. \tag{87}$$

Now since $\xi_i = v^i \cdot x$,

$$\sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \lambda_i \xi_i^2 = \sum_{i=1}^{n} \frac{(\lambda_i)^{j}}{\Delta'(\lambda_i)} x^* v^i (v^i)^* x = x \cdot Q_j x \tag{88}$$

where

$$Q_j \triangleq \sum_{i=1}^{n} \frac{(\lambda_i)^{j}}{\Delta'(\lambda_i)} v^i (v^i)^*. \tag{89}$$

Since v^i is an eigenvector of A^* [1, (122a)],

$$A^* Q_j = \sum_{i=1}^{n} \frac{(\lambda_i)^{j+1}}{\Delta'(\lambda_i)} (v^i)(v^i)^* = Q_{j+1}, \quad (j = 1, \cdots, n-1). \tag{90}$$

However, by [1, (105), (22a)]

$$Q_1 = \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} v^i (v^i)^* = V \left\{ \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} e^i (e^i)^* \right\} V^* =$$

$$= (D^{-1})^* Z \left\{ \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} e^i (e^i)^* \right\} Z^* D^{-1}. \tag{91}$$

Define the matrix E by

$$E \triangleq \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} z^i (z^i)^*, \tag{92}$$

where the z^i are the columns of Z. Then by [1, (75)], the $(\nu, \mu)^{\underline{th}}$ element of E is

$$e^\nu \cdot E e^\mu = \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} (\lambda_i)^{\nu-1} (\lambda_i)^{\mu-1}. \tag{93}$$

In Appendix 3, it is shown that

$$e^\nu \cdot E e^\mu = \sum_{i=1}^{n} \frac{(\lambda_i)^{\nu+\mu-1}}{\Delta'(\lambda_i)} = \beta_{\nu+\mu-n}, \qquad (\nu, \mu = 1, \cdots, n), \tag{94}$$

when the β's are defined as in (67a, b, c). Thus the theorem is proved for simple matrices A.

However, since (40) is analytic in a neighborhood of $x = 0$, there exist vectors $\ell^i = \ell^i(A, a)$ and matrices $R_i(A, a)$ such that

$$\sigma_j = \ell^j \cdot x - \frac{1}{2} \epsilon (x \cdot R_j x) + \cdots, \qquad (j = 1, 2, \cdots, n), \tag{95}$$

for all A. Furthermore, $\ell^i(A, a)$ and $R_i(A, a)$ are rational functions of the elements of (A, a). But the expressions in (84) are well-defined rational functions of (A, a) whether or not A is simple, and it has just been proved that

$$\ell^j = (A^*)^{j-1} b, \quad R_j = (A^*)^{j-1} Q_1, \quad (j = 1, \cdots, n), \tag{96}$$

whenever A is simple. Hence by the continuity argument used after (65), the relationships (96) must remain valid for all matrices A, simple or not. This concludes the proof.

Note: The Jacobian matrix for the transformation defined in (40) is, by (84a), $L = [b, A^* b, \cdots, (A^*)^{n-1} b]$. From [1, (16)] det L = det D and L is non-singular if the system (32) is controllable. Thus the σ_i ($i = 1, 2, \cdots, n$) defined by (40) are indeed functionally independent at $x = 0$.

CONCLUSIONS

The functions $\sigma_i(x, \epsilon)$ may be mechanized to any desired degree of accuracy by means of (45), or (66)-(67), or (84). This facilitates the synthesis of optimal feedback control systems as indicated in the discussion following Theorem 5.

APPENDIX 1

IMPLICIT FUNCTION THEOREM

Consider the m-valued vector function $f(x, y)$ where $x = (x_1, x_2, \cdots, x_n)^*$ and $y = (y_1, y_2, \cdots, y_m)^*$. Suppose that $f(x, y)$ has continuous first partial derivatives with respect to the components of y in a neighborhood of a point $(\hat{x}, \hat{y})$. If

(i) $$f(\hat{x}, \hat{y}) = 0 , \tag{1.1}$$

(ii) $$\det \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_2}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_1} \\ \frac{\partial f_1}{\partial y_2} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_m}{\partial y_2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial f_1}{\partial y_m} & \frac{\partial f_2}{\partial y_m} & \cdots & \frac{\partial f_m}{\partial y_m} \end{bmatrix}_{(\hat{x}, \hat{y})} \neq 0 , \tag{1.2}$$

then there exists in a neighborhood of $(\hat{x}, \hat{y})$ a unique set of functions $g_i = g_i(x)$, $(i = 1, 2, \cdots, m)$, such that

$$\begin{aligned} y_1 &= g_1(x) \\ y_2 &= g_2(x) \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ y_m &= g_m(x) \end{aligned} \tag{1.3}$$

represents the solution of $f(x, y) = 0$ near $(\hat{x}, \hat{y})$ in the sense that

$$f(x, g(x)) \equiv 0 \tag{1.4}$$

is valid in this neighborhood.

APPENDIX 2

Theorem. If the roots λ_i of $\Delta(s) = 0$ are distinct, then

$$\sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} = \delta_{jn}, \qquad (j = 1, 2, \cdots, n). \qquad (2.1)$$

Proof. (D. C. Lewis). By the theory of residues,

$$\oint_{|s| = \rho} \frac{s^{j-1}}{\Delta(s)} ds = 2\pi\sqrt{-1} \sum_{i=n}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)}, \qquad (j = 1, 2, 3 \cdots), \qquad (2.2)$$

where $\rho > \max |\lambda_i|$, $(i = 1, 2, \cdots, n)$. Now evaluating the above integral directly as ρ becomes arbitrarily large yields

$$\oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} ds = \lim_{\rho \to \infty} \int_0^{2\pi} \frac{2\pi \sqrt{-1}\, d\gamma}{\sum_{m=0}^{n} \alpha_m \rho^{(m-j)} \exp[(m-j)\sqrt{-1}\,\gamma]}, \qquad (2.3)$$

where $s = \rho \exp(\sqrt{-1}\,\gamma)$. For $j = 1, 2, \cdots, n-1$, then,

$$\oint_{|s| = \rho} \frac{s^{j-1}}{\Delta(s)} ds = 0, \qquad (2.4)$$

and for j = n

$$\oint_{|s| = \rho} \frac{s^{j-1}}{\Delta(s)} \, ds = \int_0^{2\pi} \sqrt{-1} \, d\gamma = 2\pi\sqrt{-1} \tag{2.5}$$

Thus from (2. 2) and (2. 5) the theorem is proven.

APPENDIX 3

<u>Theorem.</u> If the roots of $\Delta(s) = 0$ are distinct, then

$$\sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} = \beta_{j-n}, \qquad (j = 1, 2, 3, \cdots). \tag{3.1}$$

Proof. From (79)

$$\oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)}\, ds = \oint_{|s|=\rho} \sum_{\nu=0}^{\infty} \beta_\nu\, s^{-n-\nu+j-1}\, ds. \tag{3.2}$$

By the theory of residues $[-n-\nu+j = 0$ when $\nu = j-n]$

$$\oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)}\, ds = 2\pi\sqrt{-1}\, \beta_{j-n}.$$

Combining this with(2.2) gives the desired result immediately.

REFERENCES

1. Bass, R. W. and I. Gura, "Linear Canonical Forms for Controllable Systems," (to appear)

2. Bass, R. W., D. C. Lewis, et al, Final Report Contract NASr-25, Aeronca Report 388, July, 1962.

3. Lewis, D. C. and P. Mendelson, "Contributions to the Theory of Optimal Control," Trans. Amer. Math. Soc, Vol 110, No. 2, pp 232-244.

4. Hartman, P., Ordinary Differential Equations, John Wiley, New York, 1964.

5. Solodovnikov, V. V., Introduction to the Statistical Dynamics of Automatic Control, Dover, New York, 1960.

6. Bellman, R., I. Glicksberg and O. A. Gross, Some Aspects of Mathematical Theory of Control Processes, The Rand Corp. Project Rand R-313 (January, 1958).

7. Pontriagin, L. S., V. G. Boltianski, R. V. Gamkrelidze, E. F. Mischenko, The Mathematical Theory of Optimal Processes. Interscience, New York, 1962.

8. Bass, R. W. and R. F. Webber, "On Synthesis of Optimal Bang-Bang Feedback Control Systems with Quadratic Performance Index, Preprints of Papers 1965 JACC (to appear).

9. Bass, R. W. and I. Gura, "Synthesis of Time-Optimal Feedback Control Systems" (to appear).

APPENDIX C

HIGH ORDER SYSTEM DESIGN VIA STATE-SPACE CONSIDERATIONS

by

R. W. Bass and I. Gura
Space Systems Division
Hughes Aircraft Company
under

Contract No. NAS8-11421

Presented at 1965 JACC June 1965

Space Systems Division
AEROSPACE GROUP
Hughes Aircraft Company · Culver City, California

HIGH ORDER SYSTEM DESIGN VIA STATE-SPACE CONSIDERATIONS

R. W. Bass and I. Gura, Guidance and Controls Division, Aerospace Group
Hughes Aircraft Company, Culver City, California

Abstract

For n^{th} order constant plants, it is known (Letov, 1960) how to pick n desired closed-loop poles guaranteeing optimality relative to quadratic integral criteria. Also it is known (Bass, 1961) how to synthesize closed-loop poles arbitrarily by state-variable feedback, provided Kalman's criterion of controllability is satisfied. In this work these principles are combined into a unified design procedure incorporating the algorithm of Leverrier (1840). If only $m < n$ outputs can be measured, an ideal system can be synthesized "asymptotically" by a feedback filter which processes the outputs, provided Kalman's criterion of observability is satisfied. If the filter is physically realizable by a passive network, the absolute smallest number of new poles which must be introduced for mere stability is in general $[(n/m)-1]$. But the only general designs of the filter are those of Kalman (1961) and Luenberger (1964) which introduce, respectively, n and n - m new poles. Here a closed-form computer oriented general synthesis algorithm is presented which designs the filter to have only about $[(n/m) - 1)]$ poles.

Introduction: Matrix Transfer Functions and the Resolvent; Leverrier's Algorithm

Consider the open-loop system (uncontrolled system or plant) which evolves in time according to the differential equation

$$\dot{x} = Ax, \qquad x(0) = x^{o}. \tag{1}$$

Let s be a complex variable, and let $\mathcal{L}$ denote the Laplace transform operator; write x(s) for $\mathcal{L}x(t)$. Applying $\mathcal{L}$ to (1), obtain $sx(s) - x^o = Ax(s)$, or $x(s) = (sI - A)^{-1}x^o$. By Cramer's rule, the resolvent matrix $(sI-A)^{-1}$ is such that each of its elements is a ratio of polynomials in s (transfer function), and is defined whenever s is not a root of $\Delta(s) = 0$, where

$$\Delta(s) \triangleq \det(sI - A) = \sum_{i=0}^{n} \alpha_i s^i, \qquad (\alpha_n = 1) \tag{2}$$

is the open-loop characteristic polynomial. Now clearly the general solution of (1) is

$$x(t) = \exp(At)x^o, \qquad \exp(At) = \mathcal{L}^{-1}\{(sI-A)^{-1}\}, \tag{3}$$

where each element of the state-transition matrix exp (At) is the inverse Laplace transform of the corresponding element of the resolvent. A more explicit form of the resolvent can be defined ($\triangleq$) in terms of the matrix polynomial [numerator transfer matrix]

$$\Gamma(s) \triangleq \sum_{i=1}^{n} s^{i-1} S_i, \tag{4}$$

where the matrices S_i are defined for $i = 0, 1, 2, \cdots, n$ by

$$S_i \triangleq \sum_{j=i}^{n} \alpha_j A^{j-i}, \qquad (S_n = I). \tag{5}$$

Now it is well known [4] that the resolvent [open-loop transfer matrix] is

$$(sI - A)^{-1} = \Gamma(s)/\Delta(s) = \sum_{i=1}^{n} \left[\frac{s^{i-1}}{\Delta(s)}\right] S_i. \tag{6}$$

The theoretical definitions (2), (5) are useless for large n since they involve n! multiplications. Alternatively, a recursive algorithm for computing $\alpha_o, \alpha_1, \cdots, \alpha_{n-1}$ and $S_1, S_2, \cdots, S_{n-1}$ in about n^4 multiplications can be derived from Newton's identities between the α_i and the elementary symmetric functions of the roots [open-loop poles or plant poles] of (2); this is [6, 7] Leverrier's Algorithm (1840), sometimes called by other names [8] since it has been independently rediscovered or improved by Horst (1935), Souriau (1948), Frame (1949), and Faddeev and Sominskii (1949). The algorithm is, for $(j = 1, 2, \cdots, n)$,

$$\alpha_n \triangleq 1, \qquad S_n \triangleq I, \tag{7a}$$

$$\alpha_{n-j} = -(1/j)\operatorname{tr}(AS_{n-j+1}), \quad S_{n-j} = \alpha_{n-j}I + AS_{n-j+1}. \tag{7b}$$

As an automatic self-check on round-off error, note that (in theory) $S_o = 0$. The first mention of (7) in control literature appears to be that of Zadeh and Desoer [5] in 1963, although one of the present authors has used (7) in actual control design since 1960 [1, 2].

Relation between Open-loop Poles and Closed-loop Poles

Suppose that the rate of change of the state vector is modified by a forcing function $\psi_o a$, where the scalar function $\psi_o = \psi_o(x)$ is the feedback control law and the constant vector $a = (a_i)$ is the actuator vector. Thus

$$\dot{x} = Ax + a\psi_o. \tag{8}$$

For linear control

$$\psi_o = g \cdot x = g^* x, \tag{9}$$

where matrix transposition is denoted by an asterisk (*). (Vectors are columns unless otherwise specified.) Thus (8)-(9) becomes the closed-loop system

$$\dot{x} = (A + ag^*)x = \tilde{A}x \tag{10}$$

and

$$\tilde{\Delta}(s) \triangleq \det(sI - \tilde{A}) = \sum_{i=0}^{n} \tilde{\alpha}_i s^i \tag{11}$$

defines the closed-loop characteristic polynomial. Since computer algorithms for finding the roots (given the $\tilde{\alpha}_i$), or conversely, for synthesizing the $\tilde{\alpha}_i$ (given the roots), are standard, the specification of the system poles and of $\tilde{\Delta}$ will be treated as equivalent propositions. Although there are various ways of choosing a desirable $\tilde{\Delta}$, it will be assumed for the present that this choice is not an issue. Of course, it is required that $\exp(\tilde{A}t)$ decay to zero as t increases; hence the system poles must have negative real parts. Accordingly, $\tilde{\Delta}$ must be a Hurwitz polynomial. Here the relationship between $\tilde{\Delta}$ and Δ is analyzed assuming only that the gain vector g is known.

By means of the numerator transfer matrix $\Gamma(s)$ it will be shown that [1, 2]

$$\tilde{\Delta}(s) = \Delta(s) - g \cdot \Gamma(s) a, \qquad (12a)$$

or, equivalently

$$\tilde{\alpha}_i = \alpha_i - g \cdot S_{i+1} a, \qquad (i = 0, 1, \cdots, n-1). \quad (12b)$$

The proof of (12) rests on the determinantal identity

$$\det(I + cd^*) = 1 + d \cdot c. \qquad (13)$$

To establish (13), note that (because a determinant is an alternating multilinear function of its column vectors) $\det(I + cd^*) = \det(e^1 + d_1 c, \cdots, e^n + d_n c) = \det(e^1, \cdots, e^n) + d_1 \det(c, e^2, \cdots, e^n) + \cdots + d_n \det(e^1, e^2, \cdots, c) = 1 + d_1 c_1 + \cdots + d_n c_n = 1 + d \cdot c$. Now (12a) follows trivially from (6) and (13) since $\tilde{\Delta}(s) = \det(sI - A - ag^*) = \det(sI - A)\det(I - [\Gamma(s)/\Delta(s)]ag^*) = \Delta(s)\{1 - [1/\Delta(s)] g \cdot \Gamma(s) a\} = \Delta(s) - g \cdot \Gamma(s) a$. (Note that (12) is the basic lemma in Kalman's 1964 paper [3], where (13) is referred to as a "well-known matrix identity"; recently Kalman has acknowledged [1], [2] as his source. See Appendix.)

Controllability and Synthesis of Arbitrarily Specified Closed-loop Poles by State-variable Feedback

The system (8) is controllable in the sense of Kalman [11] when

$$\det D \neq 0, \quad D \triangleq (a, Aa, A^2 a, \cdots, A^{n-1} a). \qquad (14)$$

Accordingly, the system of linear algebraic equations

$$(A^{i-1} a) \cdot b = \delta_{in}, \quad (i = 1, 2, \cdots, n), \qquad (15a)$$

where δ_{in} is the Kronecker delta, or, equivalently

$$D^* b = e^n, \qquad (b = (D^{-1})^* e^n), \qquad (15b)$$

has a unique solution $b \neq 0$ if and only if the system is controllable. (e^n is the n^{th} column of the identity matrix.) The vector b is important because the system (8) is precisely equivalent [10] to the scalar n^{th} order system

$$\Delta(d/dt)\theta_1 = \psi_o \qquad (16)$$

under the explicit, reciprocal transformations

$$\theta_1 = b \cdot x, \qquad x = \{\Gamma(d/dt)\theta_1\} a. \qquad (17)$$

One may compute b from (15a) by Gaussian elimination [7], which in general requires only $(1/n)^{th}$ of the arithmetic labor of computing D^{-1}. Furthermore, once b is known, D^{-1} is known explicitly, for in [10] the present authors have established the useful matrix identities [$\det D \equiv 1/\det L$]

$$D^{-1} \triangleq (a, Aa, \cdots, A^{n-1} a)^{-1} = (S_1^* b, S_2^* b, \cdots, S_n^* b)^*, \quad (18a)$$

$$L^{-1} \triangleq (b, A^* b, \cdots, (A^*)^{n-1} b)^{-1} = (S_1 a, S_2 a, \cdots, S_n a)^*. \quad (18b)$$

The linear relations (12) may be collected into the vector equation

$$(S_1 a, S_2 a, \cdots, S_n a)^* g = -\sum_{i=1}^{n} (\tilde{\alpha}_{i-1} - \alpha_{i-1}) e^i. \qquad (19)$$

Now from (18b) and the Fredholm Alternative [15] for singular equations, the following result [1], [2] may be concluded. The system (8)-(9) may be synthesized with arbitrarily specified closed-loop poles if and only if it is controllable, in which case the gain vector g is, explicitly,

$$g = -\sum_{i=1}^{n} (\tilde{\alpha}_{i-1} - \alpha_{i-1})(A^*)^{i-1} b. \qquad (20)$$

The execution of (20) on a digital computer, via Leverrier's algorithm for finding the α_i from A, takes but a few seconds. As a self-check, the authors' program also computes $\tilde{A} = A + ag^*$ and then reapplies Leverrier's algorithm to verify that the synthesized $\tilde{\Delta}$ agrees with the specified $\tilde{\Delta}$.

Observability and Practical Asymptotic Realization of Ideal System by Lowest-order Feedback Filter

The utility of the gain vector g computed by (20) might be doubted, in that for large n not all state-variables x_i may be measured by convenient instrumentation. Typically, the only available system output is a set of m independent, known linear combinations of the x_i, say

$$y_i = h^i \cdot x, \qquad (i = 1, \cdots, m; \ 1 \leq m < n). \qquad (21a)$$

Thus the output-vector y is defined by

$$y = H^* x, \qquad H = (h^1, \cdots, h^m), \qquad (21b)$$

where the known $n \times m$ matrix H has rank $m < n$. The system is no longer defined by (8)-(9), but by (21) together with

$$\dot{x} = Ax + a\psi_o, \qquad \psi_o = \mathcal{F}(y), \qquad (22)$$

where ψ_o at time t is no longer a function of the instantaneous state x(t), but rather a "functional" (operator) on y which depends not only on y(t), but also on its past history $\{y(\tau) | 0 \leq \tau \leq t\}$.

The most precise approach to filter design is based upon Kalman's generalization [18] of the Weiner-Kolmogorov theory of optimal extraction of signals from noise. It can be proved [19], [20] that when the choice of the ideal system $\dot{x} = \tilde{A}x$ is optimized according to a quadratic performance criterion, the problem of optimal choice of ψ_o in (21)-(22) can be split into two separate problems. The first deals with optimal choice of g, and the second deals with real-time minimal-variance unbiased estimates $\hat{x}$ of x. If $y = H^* x + w$ where w is a Gaussian white noise process of a priori known spectral power density, then $\psi_o = g \cdot \hat{x}$. It can be proved {unify [12], [18] as in [19]-[20]; then specialize to the autonomous case as in [3]; finally, convert to scalar form, as in (16), by transformations analogous to (17)} that the optimal control law ψ_o can be synthesized by feeding back the observed outputs y_i through a suitable passive linear filter as in Figure 1. (The $\rho_i(s)$ are physically realizable transfer functions having the same poles but different zeros.) However, such a filter requires n poles for an n^{th} order system. Unfortunately, for large n this approach, although precise, is impractical in many applications.

Abandoning the attempt to estimate ψ_o optimally, a somewhat more economical realization theory may be developed [16] wherein the number of filter poles is equal to n-m. In this theory,

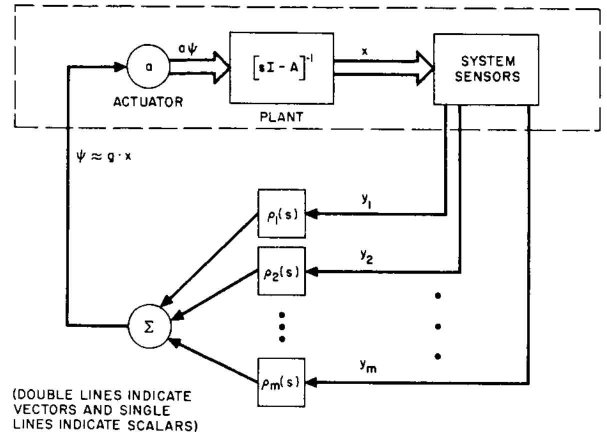


Figure 1. Closed loop system.

the ideal closed-loop system (including filter) may be specified arbitrarily.

For both theories [18] and [16], an essential hypothesis is that of plant observability, defined as

$$\operatorname{rank}[H, A^* H, (A^*)^2 H, \dots, (A^*)^{n-1} H] = n \tag{23}$$

The present theory, however, is based on ideal-system observability

$$\operatorname{rank}[H, (\tilde{A}^*) H, (\tilde{A}^*)^2 H, \dots, (\tilde{A}^*)^{n-1} H] = n \tag{24}$$

and is essentially different from the others. [Note that either (23) or (24) may hold while the other fails.] It will be shown that if $m \geq 2$, (24) yields

$$[(n/m) - 1] \leq n - \nu < n - m \tag{25}$$

where $n-\nu$ is the number of filter poles required. Often, $n-\nu$ can be arranged to equal or approximate the lower bound in (25).

Refer to the configuration of Figure 1. In terms of Laplace transforms it is clear that $\psi_o(s) = \Sigma_{i=1}^{m} \rho_i(s) y_i(s)$. Call the common poles of the $\rho_i(s)$ the open-loop filter poles, and let them be the roots of a polynomial $\Delta_{n-\nu}(s) \triangleq \Sigma_{j=0}^{n-\nu} \gamma_j s^j$. Similarly, let the zeros of $\rho_i(s)$ be the roots of $\Delta_{(i)}(s) \triangleq \Sigma_{j=0}^{\nu_i} \gamma_{ij} s^j$, $(i = 1, 2, \cdots, m)$. Then $\rho_i(s) = \Delta_{(i)}(s)/\Delta_{n-\nu}(s)$, and in the time-domain the complete system is given by (8), (21b), and

$$\Delta_{n-\nu}(d/dt)\psi_o = \sum_{i=1}^{m} \Delta_{(i)}(d/dt) y_i . \tag{26}$$

Applying the transformation (17), the system reduces to the scalar form

$$\Delta(d/dt)\theta_1 = \psi_o , \tag{27a}$$

$$\Delta_{n-\nu}(d/dt)\psi_o = \sum_{i=1}^{m} \Delta_{(i)}(d/dt) \sum_{j=1}^{n} (h^i \cdot S_j a) \theta_1^{[j-1]} \tag{27b}$$

where $\theta_1^{[i]} = d^i\theta_1/dt^i$. From (27) it can be seen that

$$\tilde{\Delta}_{2n-\nu}(d/dt)\theta_1 = 0, \tag{28a}$$

$$\tilde{\Delta}_{2n-\nu}(s) \triangleq \Delta(s)\Delta_{n-\nu}(s) - \sum_{i=1}^{m} \Delta_{(i)}(s) \sum_{j=1}^{n} (h^i \cdot S_j a) s^{j-1} \tag{28b}$$

where $\tilde{\Delta}_{2n-\nu}(s)$ is a polynomial of degree $2n-\nu$ whose roots are the actual overall system poles.

Let the open-loop filter poles $(\Delta_{n-\nu})$ as well as the ideal system poles $(\tilde{\Delta})$ be specified arbitrarily. Then the only unknowns in (28) are the polynomials $\Delta_{(i)}$, $(i = 1, 2, \cdots, m)$, whose determination completes the design of the feedback filter. For physical realizability of the filter alone, $0 \leq \nu_i \leq n-\nu$, $(i = 1, 2, \cdots, m)$ must be satisfied.

As part of the closed-loop system, the filter will be said to realize the ideal system if

$$\tilde{\Delta}_{2n-\nu}(s) = \tilde{\Delta}(s)\tilde{\Delta}_{n-\nu}(s), \tag{29}$$

where $\tilde{\Delta}_{n-\nu}$ is a Hurwitz polynomial whose roots will be called the closed-loop filter poles. The realization will be called asymptotic if the open-loop and closed-loop filter poles tend to coincide when the real parts of either set are moved uniformly toward negative infinity.

Assume the validity of (24) and seek an asymtotic realization in which $\tilde{\Delta}$ and $\Delta_{n-\nu}$ are specified arbitrarily, and the coefficients γ_{ij} of the $\Delta_{(i)}$ are determined as linear combinations of the (arbitrary) coefficients γ_i of $\Delta_{n-\nu}$. The relation between γ_{ij} and the γ_i will at first be inferred from a heuristic argument; then it will be shown that a filter designed by this method is indeed asymptotic.

Referring to (27b), attempt to choose the $\Delta_{(i)}$ so that

$$\Delta_{n-\nu}(d/dt)(g \cdot x) = \sum_{i=1}^{m} \Delta_{(i)}(d/dt)(h^i \cdot x) . \tag{30}$$

This could be true identically if $x = x(t)$ solved $\dot{x} = \tilde{A}x$ exactly; but note the transient introduced by the filter. Proceeding, however, on this "asymtotic" assumption, $d^i x/dt^i = (\tilde{A})^i x$, holds, with x^o arbitrary, and (30) reduces to

$$\Delta_{n-\nu}(\tilde{A}^*) g = \sum_{i=1}^{m} \Delta_{(i)}(\tilde{A}^*) h^i \tag{31}$$

Define $r \triangleq (\gamma_o, \gamma_1, \cdots, \gamma_{n-\nu})^*$, $Q \triangleq (g, A^* g, \cdots, (A^*)^{n-\nu} g)$, and, noting the corresponding dimensions, define matrices and vectors, for $(i = 1, \cdots, m)$, by

$$K_i \triangleq (h^i, \tilde{A}^* h^i, \cdots, (\tilde{A}^*)^{\nu_i} h^i), \qquad [n \times (\nu_i + 1)] ;$$

$$d^i \triangleq (\gamma_{i0}, \gamma_{i1}, \cdots, \gamma_{i\nu_i})^*, \qquad [1 \times (\nu_i + 1)] ;$$

$$K \triangleq (K_1, \cdots, K_m), \quad \text{and} \quad d \triangleq [(d^1)^*, \cdots, (d^m)^*]^* .$$

Then the condition (31) can be expressed as $Kd = Qr$. The smallest $n-\nu$ must satisfy $n-\nu = \max\{\nu_1, \cdots, \nu_m\}$. On the other hand, for K to be an $n \times n$ matrix, the dimensions ν_i must satisfy $n = (\nu_1 + 1) + (\nu_2 + 1) + \cdots + (\nu_m + 1)$ or $\nu_1 + \nu_2 + \cdots + \nu_m = n - m$. From this (25) follows immediately. If $\nu_1 = \nu_2 = \cdots = \nu_m$, then $\nu_i = [(n/m) - 1] = n - \nu$. Note that the columns of K can be arranged at will if the elements of d are adjusted accordingly. In particular, generate a new matrix $\hat{K}$ as follows.

(a) Start with the columns $h^1, h^2, \cdots, h^m$.

(b) Adjoin to this the columns $(\tilde{A}^*)h^1, (\tilde{A}^*)h^2, \cdots, (\tilde{A}^*)h^m$ one by one, checking that each new column is linearly independent of the previous ones. (Use the Gram-Schmidt orthogonalization procedure.)

(c) If any of the new columns is found to be dependent, omit it from the matrix and go on to the next.

(d) Continue adjoining columns until n linearly independent ones have been found.

(e) After $(\tilde{A}^*)h^m$ has been tested, continue with $(\tilde{A}^*)^2 h^1, \cdots, (\tilde{A}^*)^2 h^m$, $(\tilde{A}^*)^3 h^1, \cdots, (\tilde{A}^*)^3 h^m, \cdots$, $(\tilde{A}^*)^{n-1} h^1, \cdots, (\tilde{A}^*)^{n-1} h^m$.

(f) If a column $(\tilde{A}^*)^i h^j$ has been skipped because of linear dependence, all columns of the form $(\tilde{A}^*)^{\ell} h^j$ where $\ell > i$ can be skipped, immediately, since they also must be dependent on the previous columns.

When observability applies in the form (24), there must be n linearly independent columns in the matrix $\hat{K}$ generated as described above. If $\hat{d}$ is the correspondingly ordered vector of n unknowns, the system $\hat{K}\hat{d} = Qr$ can be solved for the elements of $\hat{d}$, which (after appropriate re-ordering) define the numerator polynomials $\Delta_{(i)}$ of the desired feedback filter.

Previously, it was required that the filter in question obey (29). This can be verified by an algebraic manipulation which is both tedious and rather subtle. Indeed, it can be shown that $\Delta_{n-\nu}$ and $\tilde{\Delta}_{n-\nu}$ are related by the equation

$$\tilde{\gamma}_{i-1} = \gamma_{i-1} + \sum_{k=i}^{n-\nu} \gamma_k (g \cdot A^{k-i} a) - \sum_{\ell=1}^{m} \gamma_{\ell k}(h^{\ell} \cdot A^{k-i} a) \qquad (32)$$

for $(i = 1, 2, \cdots, n-\nu)$, where now one defines $\gamma_{\ell k} = 0$ for $k > \nu_{\ell}$. From the form of (32) the asymptotic equivalence between $\Delta_{n-\nu}$ and $\tilde{\Delta}_{n-\nu}$ follows readily.

Optimal Choice of Closed-loop Poles

Up to this point it has been shown that, given a desired $\tilde{\Delta}(s)$, a unique gain vector g can be found (20) so that the ideal system

$$\dot{x} = \tilde{A}x \equiv (A + ag^*)x = Ax + a\psi_o, \quad \psi_o = g \cdot x, \quad x(0) = x^o \qquad (33)$$

is synthesized by $\psi_o = g \cdot x$. When only an output $y = H^* x$ is observed, ψ_o can be asymptotically synthesized as $\psi_o = \mathcal{F}(y)$ by means of the feedback filter (26).

For large n, however, the available arbitrariness in specification of $\tilde{\Delta}(s)$ constitutes an "embarrassment of riches." To remedy this, the question of choosing a control law that will optimize some performance criterion will be considered. For present purposes define this criterion as

$$\Phi = \tfrac{1}{2}\int_0^{\infty} (x \cdot Cx + \psi_o^2)\, dt, \qquad C = C^* \geq 0, \qquad (34)$$

and call the control ψ_o "optimal" if it minimizes (34).

The choice of an appropriate matrix C is important. It must be done in the context of a specific problem. For example, in aerospace vehicle stabilization it may be required to maintain certain quantities $|q^i \cdot x|$, $(i = 1, 2, \cdots, \tilde{m})$, such as "structural load", "pitch error", etc., below stated bounds while minimizing the maximum over future time of some critical quantity $|q^{\tilde{m}} \cdot x|$, such as "lateral drift." The important minimax control problem may be solved to a first approximation [exact solution requires nonlinear feedback] by noting that, in the integral $\int_0^{\infty} (q^{\tilde{m}} \cdot x(t)/\kappa_o)^2 dt$, the total contribution of times at which $|q^{\tilde{m}} \cdot x(t)| > \kappa_o$ holds is "penalized" disproportionately in comparison to that of times at which $|q^{\tilde{m}} \cdot x(t)| < \kappa_o$ holds. Hence it would be desirable to find a performance criterion which minimizes the above integral while at the same time maintaining $\int_0^{+\infty} (q^i \cdot x)^2 dt$, $(i = 1, 2, \cdots, \tilde{m}-1)$ and $\int_0^{+\infty} \psi_o^2 dt$ within required bounds. All this can be accomplished by defining

$$C \triangleq \kappa_1 q^1 (q^1)^* + \cdots + \kappa_{\tilde{m}} q^{\tilde{m}} (q^{\tilde{m}})^* \qquad (35)$$

in (34). If $\tilde{m} \geq n$, and the q^i are linearly independent, C is positive definite. The theoretical development of this case is more straightforward than that for which $\tilde{m} < n$ and C is only guaranteed non-negative definite.

Another approach to choosing C can be found in the "implicit model reference" method mentioned by Soviet authors such as Aizerman. Basically, it is desired to force $\theta_1 = b \cdot x$ to behave in the mean like solutions of $\Delta_{\hat{n}}(d/dt)\theta_1 = 0$ where $\Delta_{\hat{n}}(s)$ is a Hurwitz polynomial of degree $\hat{n} < n$. Using (17) it is clear that $\Delta_{\hat{n}}(d/dt)\theta_1 = [\Delta_{\hat{n}}(A^*)b] \cdot x$ whence the matrix C to be used in (34) is

$$C \triangleq qq^*, \qquad q \triangleq \Delta_{\hat{n}}(A^*)b. \qquad (36)$$

The first general results on the solution of the problem described in (33)-(34) are due to Bellman, Glicksberg and Gross in 1954 (cf. [9]). After slight modification of their derivation, it can be shown that their work gives, for $C > 0$

$$\psi_o = a \cdot p, \qquad (37)$$

where the "co-state" vector p satisfies the two-point Lagrangian boundary-value problem defined by (1) and

$$\ddot{p} + (A^* - CAC^{-1})\dot{p} - C(AC^{-1}A^* + aa^*)p = 0, \quad x(+\infty) = 0. \qquad (38)$$

However, the numerical methods they suggest for solving (38) apply only for fixed x^o and do not yield (37) in the feedback form $\psi_o(x)$ needed for synthesis.

In 1960, Letov [21] implicitly assumed controllability via the use of Lur'e coordinates [10] and improved (37) by showing that under slight restrictions there exists a constant matrix B such that

$$p = -Bx, \qquad (39)$$

whence the optimal control law determined by (34) is linear:

$$\psi_o = g \cdot x, \qquad g \equiv -Ba. \qquad (40)$$

Letov applied the classical Euler-Lagrange necessary conditions to (33)-(34), and expressed the result in Hamiltonian form [readily seen equivalent to (38)]

$$\begin{pmatrix}\dot{x}\\ \dot{p}\end{pmatrix} = \mathbf{H}\begin{pmatrix}x\\ p\end{pmatrix}, \qquad \mathbf{H} \triangleq \begin{pmatrix}A, & aa^*\\ C, & -A^*\end{pmatrix}. \tag{41}$$

After defining $\Delta_{2n}(s) \triangleq \det(s\mathbf{I}_{2n} - \mathbf{H})$ and showing that $\Delta_{2n}(s)$ is a polynomial in even powers of s only, he concluded that if the roots of this polynomial are distinct and non-imaginary, the n Hurwitz roots are the optimal poles of (33)-(34). Hence

$$\Delta_{2n}(s) = (-1)^n\tilde{\Delta}(-s)\tilde{\Delta}(s). \tag{42}$$

An explicit expression for $\Delta_{2n}(s)$ can be obtained in the following way:

Define

$$\mathbf{K} = \begin{bmatrix} I, & (sI-A)^{-1}aa^* \\ 0, & I \end{bmatrix}, \tag{43}$$

where $\det \mathbf{K} = 1$, and argue that

$$\begin{aligned}\Delta_{2n} &= \det[(s\mathbf{I}_{2n} - \mathbf{H})\mathbf{K}] = \\ &= \det(sI-A)\det[(sI+A^*) - C(sI-A)^{-1}aa^*] = \\ &= \Delta(s)\det(sI+A^*)\det[I-(sI+A^*)^{-1}C(sI-A)^{-1}aa^*] = \\ &= (-1)^n\Delta(s)\Delta(-s)\det[I-\{\Gamma^*(-s)/\Delta(-s)\}C\{\Gamma(s)/\Delta(s)\}aa^*]\end{aligned}$$

Application of (13) then immediately yields the desired result

$$\Delta_{2n}(s) = (-1)^n[\Delta(s)\Delta(-s) + a\cdot\Gamma^*(-s)C\Gamma(s)a]. \tag{44}$$

The results (39), (40), and (42), (44) actually apply when the roots of $\Delta_{2n}(s)$ are non-distinct and when C is only non-negative definite, provided that $x \cdot Cx$ is the square of an "observable" quantity. This can be deduced from Kalman's nearly definitive studies [12], [3], which combine Pontriagin's <u>necessary</u> Maximum Principle [17] with the <u>sufficient</u> Hamilton-Jacobi partial differential equation [9]. Kalman shows that the optimal control law for $\dot{x} = Ax + a(g\cdot x)$ defined by the criterion (34) is given by $g = -Ba$, and is stable if there exists a symmetric $B > 0$ satisfying

$$BA + A^*B - Baa^*B = -C. \tag{45}$$

Under these conditions, the function $x \cdot Bx$ is a Liapunov function for the closed-loop system $\dot{x} = \tilde{A}x$; and B must be given [9] by

$$B = \int_0^\infty \exp(\tilde{A}^*t)[C + gg^*] \exp(\tilde{A}t)\,dt. \tag{46}$$

Furthermore, $\Phi = \Phi(x^o) = \frac{1}{2}x^o \cdot Bx^o$ and $p = -\text{grad}\,\Phi = -Bx$ satisfies the necessary condition $\max_{\psi_o} \mathcal{H}(x, p, \psi_o) = 0$ where

$$\mathcal{H} \triangleq p\cdot(Ax + a\psi_o) - (\tfrac{1}{2})(x\cdot Cx + \psi_o^2). \tag{47}$$

Although Kalman suggests a method for finding g explicitly (integration of a matrix-type Riccati differential equation), a more efficient approach, for (33), can be obtained by combining his work with that of Letov and (20) above.

The results (45)-(46) are equivalent to (41), (44) as can be shown by the following argument. Rewrite (45) as $B(sI-A) - (sI+A^*)B = C - Baa^*B$. Premultiplication by $-a^*\Gamma^*(-s)/\Delta(-s)$ and postmultiplication by $(\Gamma(s)/\Delta(s))a$ yields (after multiplication by $-\Delta(-s)\Delta(s)$ and substitution on the left of $g = -Ba$) the result $-a^*\Gamma^*(-s)g\Delta(s) - g^*\Gamma(s)a\Delta(-s) = a^*\Gamma^*(-s)(C - Baa^*B)\Gamma(s)a$. Now adding $\Delta(s)\Delta(-s)$ to both sides and rearranging gives, after use of (12),

$$\tilde{\Delta}(s)\tilde{\Delta}(-s) = \Delta(-s)\Delta(s) + a\cdot\Gamma^*(-s)C\Gamma(s)a \tag{48}$$

which is exactly equivalent to (42), (44).

Note that since the $\tilde{\Delta}(s)$ determined from (48) by construction must be Hurwitz, (46) must yield a $B > 0$ if $C > 0$ and so (48) is totally consistent with (41). For a semi-definite C, additional conditions must be satisfied to ensure that $B > 0$. Consider the polynomial $a\cdot\Gamma^*(-s)C\Gamma(s)a$. Then for $C \geq 0$, $a\cdot\Gamma^*(-s)C\Gamma(s)a = \Delta_{\hat{n}}(-s)\Delta_{\hat{n}}(s)$ where $\Delta_{\hat{n}}(s)$ is a polynomial of degree $\hat{n} \leq n-1$ ($<n$) whose coefficients μ_i are readily computable. Now define a vector q by the relation $\Delta_{\hat{n}}(s) = \Sigma_{i=1}^n \mu_i s^{i-1} \triangleq q\cdot\Gamma(s)a$. Then $\mu_i = q\cdot S_i a$, $(i = 1, 2, \cdots, n)$, or $(S_1 a, \cdots, S_2 a)^* q = \hat{q}$ where $\hat{q}$ denotes the vector of coefficients of $\Delta_{\hat{n}}(s)$. Applying (18b), it is clear that $q = \Delta_{\hat{n}}(A^*)b$. Thus

$$a\cdot\Gamma^*(-s)C\Gamma(s)a \equiv [q\cdot\Gamma(-s)a][q\cdot\Gamma(s)a]. \tag{49}$$

This important result shows that, by (48), the matrix C can be replaced in Φ by a new matrix defined as in (36) without affecting the determination of the optimal control law. In other words, minimizing the Φ of (34) is exactly <u>equivalent</u> to minimizing the simpler performance criterion

$$\Phi' = \tfrac{1}{2}\int_0^{+\infty}[(q\cdot x)^2 + (g\cdot x)^2]\,dt. \tag{50}$$

Then in (46), $B > 0$ unless for some $x^o \neq 0$, $(g\cdot x) \equiv 0$ and $(q\cdot x) \equiv 0$, $0 \leq t < +\infty$. But in this case, $x(t) = \exp(\tilde{A}t)x^o \equiv \exp(At)x^o$ whence the derivatives of $q\cdot x$ at $t = 0$ become $[(A^*)^{i-1}q]\cdot x^o$, $(i = 1, \cdots, n)$. Now assume that $q\cdot x$ is <u>observable</u>; that is

$$\det[q, A^*q, \cdots, (A^*)^{n-1}q] \neq 0. \tag{51}$$

Then $q\cdot x \equiv 0$ implies the contradiction that $x^o = 0$, i.e., $q\cdot x \equiv 0$ is impossible and $B > 0$.

As an alternative to (51) in checking that $q\cdot x \not\equiv 0$ when $g\cdot x \equiv 0$, consider the following argument.

By (12a) and (17), if $\tilde{\Delta}(d/dt)\theta_1 = 0$, then

$$g\cdot x \equiv 0 \Rightarrow [g\cdot\Gamma(d/dt)a]\theta_1 \Rightarrow \Delta(d/dt)\theta_1 = 0. \tag{52a}$$

Similarly, from (36) and (52a), when $g\cdot x \equiv 0$ is satisfied,

$$q\cdot x \equiv 0 \Rightarrow \Delta_{\hat{n}}(A^*)b\cdot x = \Delta_{\hat{n}}(d/dt)\theta_1 = 0. \tag{52b}$$

Now, if $\Delta_{\hat{n}}(s)$ is a factor of $\Delta(s)$, $\Delta_{\hat{n}}(d/dt)\theta_1 = 0$ implies that $\Delta(d/dt)\theta_1 = 0$, in which case (52b) holds. Conversely, if the conditions of (52) are both satisfied, $\Delta_{\hat{n}}(s)$ must be a factor of $\Delta(s)$. This can be shown directly if the eigenvalues $\{\hat{\lambda}_i\}$ of $\Delta_{\hat{n}}(s)$ and those $\{\lambda_i\}$ of $\Delta(s)$ are distinct. Making this assumption, and recalling that $\Delta_{\hat{n}}(s) = \Pi_{i=1}^{\hat{n}}[s - \hat{\lambda}_i]$, the general solution of $\Delta_{\hat{n}}(d/dt)\theta_1 = 0$ is a linear combination of the solutions of $[(d/dt) - \hat{\lambda}_i]\theta_1 = 0$, while a similar conclusion holds regarding the solution of $\Delta(d/dt)\theta_1 = 0$ and those of $[(d/dt) - \lambda_i]\theta_1 = 0$. This implies that the $\hat{\lambda}_i, (i = 1, 2, \cdots, \hat{n})$ are included in the λ_i, i.e., $\Delta_{\hat{n}}(s)$ must be a factor of $\Delta(s)$.

In order to extend this result to the general case, define $\alpha \triangleq (\alpha_o, \alpha_1, \cdots, \alpha_{n-1})^*$. Then the companion matrix $\hat{C}$ of $\hat{A}$ or Δ is defined as usual by $\hat{C} = (e^2, \cdots, e^n, -\alpha)^*$. Now it can be shown (cf. techniques of [10]) that, referring to (18b),

$$[q, A^*q, \cdots, (A^*)^{n-1}q] \equiv L[\hat{q}, \hat{C}^*\hat{q}, \cdots, (\hat{C}^*)^{n-1}\hat{q}], \quad (\det L = 1/\det D \neq 0), \quad (53)$$

where $\hat{q}$ is as defined after (48). By controllability and (18b), the observability condition (51) is equivalent to

$$\det[\hat{q}, \hat{C}^*\hat{q}, \cdots, (\hat{C}^*)^{n-1}\hat{q}] \neq 0. \quad (54)$$

Thus (54) is now sufficient to show that $\Delta_{\hat{n}}(s)$ and $\Delta(s)$ have no common factors for distinct roots of these polynomials. However $\det[\hat{q}, \hat{C}^*q, \cdots, (\hat{C}^*)^{n-1}\hat{q}]$ is a multinomial in μ_i and α_i only; thus it must be the "resultant" (the general condition for two polynomials to have no common factors) and so (51) is satisfied if and only if $\Delta_{\hat{n}}(s)$ is not a factor of $\Delta(s)$.

It remains only to develop criteria for $\tilde{\Delta}$ to be Hurwitz, which are based on lecture notes distributed by W. M. Wonham at Purdue University. [In these notes, Wonham overlooks the necessity of a test of observability of $q \cdot x$.] If the open-loop characteristic polynomial $\Delta(s)$ has no purely imaginary roots, (48) can be written as

$$|\tilde{\Delta}(j\omega)|^2 = |\Delta(j\omega)|^2 + |q \cdot \Gamma(j\omega)a|^2 \geq |\Delta(j\omega)|^2 > 0, \quad -\infty < \omega < +\infty, \quad j = \sqrt{-1}.$$

This guarantees that none of the roots of $\Delta_{2n}(s)$ are imaginary; hence $\tilde{\Delta}(s)$ will be Hurwitz. Furthermore, if

$$S_1 a \cdot C S_1 a > 0 \quad (55)$$

is satisfied, an open-loop pole at $\omega = 0$ does not prevent $\tilde{\Delta}(s)$ from being Hurwitz since

$$|\tilde{\Delta}(0)|^2 \geq 0 + |q \cdot \Gamma(0)a|^2 = (S_1 a) \cdot C(S_1 a) > 0. \quad (56)$$

Thus if the open-loop system has no imaginary poles except possibly at $s = 0$, in which case (55) is assumed to be satisfied, $\tilde{\Delta}(s)$ must be Hurwitz.

These concepts are now unified into an actual design procedure.

(a) Choose an appropriate matrix C by the methods of (35) or (36) above.

(b) Compute $\Delta(s)$ by Leverrier's algorithm (7).

(c) Find the roots of $\Delta(s) = 0$. If $\Delta(s)$ has purely imaginary roots (other than $s = 0$) modify A until it has none. If $\Delta(0) = 0$, also check the condition

$$(S_1 a) \cdot C(S_1 a) > 0;$$

if it fails, modify C until it holds.

(d) As explained above (49), compute the vector $q = \Delta_{\hat{n}}(A^*)b$ such that

$$a \cdot \Gamma^*(-s) C \Gamma(s) a \equiv [q \cdot \Gamma(-s)a][q \cdot \Gamma(s)a] = \triangleq \Delta_{\hat{n}}(-s)\Delta_{\hat{n}}(s).$$

(e) Find the roots of $\Delta_{\hat{n}}(s) = 0$. If $\Delta_{\hat{n}}(s)$ is a factor of $\Delta(s)$, modify C until it is not.

(f) Compute the polynomial $\Delta_{2n}(s)$ by the following explicit expansion of (44):

$$\Delta(-s)\Delta(s) = \alpha_0^2 + \sum_{i=1}^{n-1} \hat{\alpha}_i s^{2i} + (-1)^n \alpha_n^2 s^{2n}, \quad (57a)$$

$$\hat{\alpha}_i \triangleq (-1)^i \alpha_i^2 + 2 \sum_{\substack{j= \\ \max \\ (0,2i-n)}}^{i-1} (-1)^j \alpha_j \alpha_{2i-j}, \quad (i = 1, \cdots, n-1), \quad (n \geq 2) \quad (57b)$$

$$a \cdot \Gamma^*(-s) C \Gamma(s) a = (S_1 a) \cdot C(S_1 a) + \sum_{i=1}^{n-2} \epsilon_i s^{2i} + (-1)^{n-1}(a \cdot Ca)s^{2n-2}, \quad (57c)$$

$$\epsilon_i \triangleq (-1)^i (S_{i+1}a) \cdot C(S_{i+1}a) + 2 \sum_{\substack{j= \\ \max \\ (0,2i-n+1)}}^{i-1} (-1)^j (S_{j+1}a) \cdot C(S_{2i-j+1}a), \quad (i = 1, \cdots, n-2), \quad (n \geq 3) \quad (57d)$$

(g) Alternatively, Leverrier's algorithm can be applied to the $2n \times 2n$ matrix $\mathbf{H}$ of (41) to give $\Delta_{2n}(s)$.

(h) Find the roots of $\Delta_{2n}(s)$, and from the n roots that have negative real parts generate the unique polynomial $\Delta(s)$ such that $\Delta_{2n}(s) = (-1)^n \tilde{\Delta}(-s)\tilde{\Delta}(s)$.

(i) Insert the coefficients of $\tilde{\Delta}$ into (20) to find the desired optimal gain vector g. In practice it is useful to compute a one-parameter family of gain vectors, say $g = g(\mu_0)$, by replacing C with $\mu_0 C$, $0 < \mu_0 < +\infty$, and letting μ_0 vary over the positive real numbers.

Intrinsic Adaptivity to Actuator Saturation

In engineering practice, of course, actuators are linear only over a finite range and have limited amplitude. By renormalizing $\|a\|$ if necessary, it can be assumed without loss of generality that, in (8),

$$|\psi_0| \leq \rho_0. \quad (58)$$

Hence it is of great interest to study the behavior of (8) under the control law

$$\tilde{\psi}_0 \triangleq \rho_0 \operatorname{sat}[\mu_0 (g \cdot x)/\rho_0], \quad (59)$$

$$\tfrac{1}{2} < \rho_0 < +\infty, \qquad \tfrac{1}{2} < \mu_0 < +\infty, \quad (60)$$

where $\operatorname{sat}[\theta_0] \triangleq \theta_0$ for $|\theta_0| \leq 1$ and $\operatorname{sat}[\theta_0] = \operatorname{sgn}[\theta_0]$ for $|\theta_0| \geq 1$. Note that if instead of (60) one requires

$$1 \leq \rho_0 < +\infty, \qquad 1 \leq \mu_0 < +\infty, \quad (61)$$

then there is a clear physical interpretation to (59) and (61); in fact,

$$\tilde{\psi}_0 = \mu_0 (g \cdot x), \quad |g \cdot x| \leq (\rho_0/\mu_0), \quad (62)$$

i.e., the control law (59) is linear at least in the region

$$\|x\| \le (\rho_o/\mu_o \|g\|), \qquad (63)$$

while increasing $\mu_o > 1$ is the same as increasing the control signal gain $\|g\|$, and increasing $\rho_o > 1$ is the same as increasing the actuator amplitude $\|a\|$.

Use of (59) permits what seems to be the first unified theory of linear, linear-saturating, and bang-bang control. Clearly the extremes are

$$\mu_o = 1, \quad \rho_o = +\infty, \ (\text{LINEAR CONTROL}, \tilde{\Psi}_o = g \cdot x) \qquad (64)$$

$$\rho_o = 1, \quad \mu_o = +\infty, \ (\text{BANG-BANG CONTROL}, \ \tilde{\Psi}_o = \operatorname{sgn}[g \cdot x]) \qquad (65)$$

There are two important performance criteria applicable to (8) and (59), namely

$$\Phi_o = \int_0^{+\infty} [x \cdot Cx + (g \cdot x)^2] dt, \qquad (66)$$

and the largest number $\lambda_o > 0$ such that for some $\gamma_o \ge 1$,

$$\|x(t)\| \le \|x^o\| \gamma_o \exp(-\lambda_o t), \ (0 \le t < +\infty), \qquad (67)$$

whenever

$$\|x^o\| \le (\rho_o/\gamma_o \|g\|). \qquad (68)$$

Referring to (45), and using $x \cdot Bx$ as a Liapunov function, it can be shown that if $g = -Ba$ is computed as in the procedure above, the system (8), (59), (68) is asymptotically stable on (60), $\|x^o\| < 2\rho_o/(\gamma_o \|g\|)$.

Moreover, neither performance criterion Φ_o or λ_o is degraded by allowing ρ_o, μ_o to vary on (61)!

This truly remarkable property of the gain vector $g = g(a, A, C)$ obviously enhances the practical usefulness of the design procedures developed above.

Conclusions

A unified practical algorithm for the design of lowest-order [physically realizable] asymptotic realizations of ideal optimal control systems is obtained by combining the just-listed procedure (a)-(i) with the procedure (a) - (f) preceding (32). The authors have implemented this in a digital computer program. Inputs to the computer are plant matrix A; actuator vector a; sensor vectors $h^1, h^2, \cdots, h^m$; performance vectors $q^1, q^2, \cdots, q^{\tilde{m}}$; trade-off coefficients $\kappa_1, \kappa_2, \cdots, \kappa_{\tilde{m}}$; and filter poles $\Delta_{n-\nu}(s)$. Outputs are optimal filter zeros $\Delta_{(i)}(s)$, $(i=1, \cdots, m)$ for asymptotic realization of the system which (in the mean-square) minimizes the performance index $\kappa_1(q^1 \cdot x)^2 + \kappa_2(q^2 \cdot x)^2 + \cdots + \kappa_{\tilde{m}}(q^{\tilde{m}} \cdot x)^2$.

Appendix

R. E. Kalman has stated that he learned (13) from [1] [2], but subsequently encountered instances of its use by Caratheodory without comment (cf. [22], p. 342). Kalman has kindly supplied the following proof, which is amusing, but technically less elementary than that given here. Since there must exist (n-1) linearly independent vectors u^i orthogonal to d, $(I + cd^*)u^i = u^i + (d \cdot u^i)c = u^i$; hence $I + cd^*$ has n-l eigenvalues $\lambda_i = 1$, $(i = 1, 2, \cdots, n-1)$. Now

$\lambda_1 + \cdots + \lambda_{n-1} + \lambda_n = (n-1) + \lambda_n \equiv \operatorname{tr}(I + cd^*) = n + d \cdot c$, whence $\det(I + cd^*) \equiv \lambda_1 \lambda_2 \cdots \lambda_{n-1} \lambda_n = \lambda_n = 1 + d \cdot c$,

Notational Conventions

a. Matrices are upper case letters.

b. Vectors are lower case unsubscripted or superscripted letters. Scalar product is $\cdot$; $\|x\|^2 \triangleq x \cdot x$.

c. Scalars are subscripted lower case letters.

d. Exceptions to these rules are i, j, k, l, ν, n, which are used as summation indices or scalars; s which is a complex scalar; $\Delta(s)$ which is a polynomial in s; and t which denotes time. Also Φ and $\mathcal{H}$ are scalars.

e. The i^{th} column of the identity matrix is represented by e^i.

f. The symbol $\triangleq$ denotes equality by definition.

g. The symbol $\equiv$ denotes identity.

References

1. R.W. Bass, Control Synthesis and Optimization, NASA Langley Research Center, August 1961 (multilithed lecture notes).
2. R.W. Bass, P. Mendelson, D.C. Lewis, et al. Aspects of General Control Theory, Final Report AFOSR 2754, Contract AF 49(638)-1003 August 1962.
3. R.E. Kalman, "When Is a Linear Control System Optimal?" Trans. ASME Journal of Basic Engineering, pp. 51-60, Ser. D, March 1964.
4. R.A. Frazer, W.J. Duncan and A.R. Collar, Elementary Matrices, Cambridge University Press. The Macmillan Co., New York, 1947.
5. L.A. Zadeh and C.A. Desoer, Linear System Theory, McGraw-Hill, New York, 1963.
6. V.N. Faddeeva, Computational Methods of Linear Algebra, Dover Publications Inc., New York, 1959.
7. A.S. Householder, Principles of Numerical Analysis, McGraw-Hill Co., New York, 1963.
8. G.E. Forsythe and L.W. Straus, "The Souriau-Frame Characteristic Equation Algorithm on a Digital Computer", J. Math. Phys., 34, pp. 152-156, 1955.
9. R. Bellman, I. Glicksberg and O.A. Gross, Some Aspects of the Mathematical Theory of Control Processes, The Rand Corp., Project Rand R-313 (January, 1958).
10. R.W. Bass and I. Gura, "Linear Canonical Forms for Controllable Systems," (to appear).
11. R.E. Kalman, "On the General Theory of Control Systems," Proc. of First IFAC Congress, Moscow, 1960.
12. R.E. Kalman, "Contributions to the Theory of Optimal Control," Bol. Soc. Mat. Mexicana, 1960, pp. 102-119.
13. R.E. Kalman, Y.C. Ho, K.S. Narendra, "Controllability of Linear Dynamical Systems," Contributions to Differential Equations, Vol. 1, Interscience, New York, 1963.

14. R.E. Kalman, "Mathematical Description of Linear Dynamical Systems," J. S.I.A.M. Control, Ser. A., Vol. 1, No. 2, pp. 153-92, 1963.

15. B. Friedman, Principles and Techniques of Applied Mathematics, John Wiley and Sons, Inc., New York, 1956.

16. D.G. Luenberger, "Observing the State of a Linear System," IEEE Trans. on Mil. Elect., April, 1964.

17. L.S. Pontriagin, V.G. Boltianskii, R.V. Gamkrelidze, E.F. Mischenko, "The Mathematical Theory of Optimal Processes," Interscience, New York, 1962.

18. R.E. Kalman, "A New Approach to Linear Filtering and Prediction Theory," J. of Basic Engr., Trans. ASME, Ser. D, Vol. 82, pp. 35-45, March, 1960.

19. J.E. Potter, "A Guidance-Navigation Separation Theorem," AIAA Paper No. 64-653, August, 1964.

20. F. Tung, "Linear Control Theory Applied to Interplanetary Guidance," IEEE Trans. on Automatic Control, January, 1964.

21. A.M. Letov, "Analytical Controller Design I," Avtomatika i Telemekhanika, Vol. 21, No. 4, pp. 436-441, April, 1960.

22. Frank-Mises, Die Differential-und Integralgleichungen der Mechanik und Physik, Brannschweig, 1930.

SUPPLEMENTARY NOTES ON ADAPTIVITY ANALYSIS

SUPPLEMENTARY NOTES ON ADAPTIVITY ANALYSIS

Assume that the control law for

$$\dot{x} = Ax + a\psi \tag{1}$$

has been formed by the optimal methods of the previous section. The effects of increases or saturation in the feedback signal or in the actuators are considered here. In general, let

$$\psi = \rho \ \mathrm{sat}\left(\frac{\mu g \cdot x}{\rho}\right). \tag{2}$$

If $|g \cdot x| \leq \rho/\mu$, or if $\rho \to \infty$

$$\psi \approx \mu g \cdot x . \tag{3a}$$

If $|g \cdot x| \geq \rho/\mu$ or if $\mu \to \infty$

$$\psi \approx \rho \ \mathrm{sgn}\,(g \cdot x), \tag{3b}$$

and so the possibility of linear and bang-bang control are included in (2).

EFFECT ON STABILITY

Asymptotic stability for a system

$$\dot{x} = f(x) \tag{4}$$

is ensured in the domain determined by $V(x) < \epsilon$ where $V(x)$ is a Liapunov function for (4) and ϵ is a positive constant if

$$V(x) < \epsilon \qquad \dot{V}(x) < 0. \tag{5}$$

Now let

$$V = \frac{1}{2} x \cdot Bx \tag{6}$$

be a Liapunov function for (1). Then

$$\dot{V} = \frac{1}{2} (Ax + a\psi) \cdot Bx + \frac{1}{2} x \cdot B (Ax + a\psi) \tag{7}$$

$$= \frac{1}{2} x \cdot (A^{*} B + BA) x + x \cdot Ba\psi .$$

Apply (2) and

$$-C = BA + A^{*}B - gg^{*}, \tag{8a}$$

$$g = -Ba \tag{8b}$$

to (7) and obtain

$$\dot{V} = -\frac{1}{2} x \cdot Cx + \frac{1}{2} (g \cdot x)^2 - \rho (g \cdot x) \text{ sat } \left(\frac{\mu g \cdot x}{\rho}\right) \tag{9a}$$

or

$$\dot{V} = -\frac{1}{2} x \cdot Cx - \frac{1}{2} |g \cdot x| \left\{ 2\rho \text{ sat } \left(\frac{\mu g \cdot x}{\rho}\right) - |g \cdot x| \right\} . \tag{9b}$$

The control ψ is linear when

$$\frac{\mu}{\rho} |g \cdot x| < 1.$$

Then (9b) gives

$$\dot{V} = -\frac{1}{2} x \cdot Cx - \frac{1}{2} |g \cdot x| \left\{ (2\mu - 1) |g \cdot x| \right\} ,$$

and stability is ensured if

$$\mu > \frac{1}{2} . \tag{10}$$

Similarly, the control is saturated when

$$\frac{\mu}{\rho} \left| g \cdot x \right| > 1. \tag{11}$$

Then

$$\dot{V} = -\frac{1}{2} x \cdot Cx - \frac{1}{2} \left| g \cdot x \right| \left\{ 2\rho - \left| g \cdot x \right| \right\}, \tag{12}$$

and

$$\left| g \cdot x \right| < 2\rho \tag{13}$$

is necessary for asymptotic stability.

To establish sufficient conditions for a region of asymptotic stability in this case, consider the lemma

$$(x \cdot Ba)^2 \leq (x \cdot Bx)\,(a \cdot Ba), \tag{14a}$$

where

$$B = B^* > 0, \tag{14b}$$

$$a \neq 0. \tag{14c}$$

Proof: If x and a are linearly dependent, the equality sign obviously holds. Alternatively, if x and a are linearly independent, let

$$u \triangleq (x \cdot Ba / a \cdot Ba)a, \qquad v \triangleq x - u \neq 0. \tag{15}$$

By direct substitution

$$v \cdot Bu = u \cdot Bv = 0, \tag{16}$$

and so

$$x \cdot Bx = (v + u) \cdot B(v + u) = v \cdot B\, v + u \cdot Bu > u \cdot Bu. \tag{17}$$

Thus by (15)

$$x \cdot Bx > (x \cdot Ba)^2/(a \cdot Ba), \tag{18}$$

and the lemma is proved.

Now by (8b) and this lemma,

$$|g \cdot x|^2 \leq (x \cdot Bx)(a \cdot Ba). \tag{19}$$

Then if

$$(x \cdot Bx)(a \cdot Ba) < (2\rho)^2, \tag{20}$$

$$V = \frac{1}{2} x^o \cdot Bx^o < \frac{2\rho^2}{a \cdot Ba} \tag{21}$$

must imply that

$$\dot{V} < 0, \tag{22}$$

and asympotic stability of (1) with saturated control will be guaranteed in the region where

$$x^o \cdot Bx^o < \frac{4\rho^2}{a \cdot Ba}, \tag{23a}$$

or equivalently where

$$x \cdot Bx^o < \frac{-4\rho^2}{a \cdot g}. \tag{23b}$$

(See Table I at the end of this appendix.)

EFFECT ON PERFORMANCE INDEX

$$\Phi(x^o) = \frac{1}{2} \int_0^\infty (x \cdot Cx + (g \cdot x)^2) \, dt \tag{24}$$

Since from (9a)

$$- \dot{V} = \frac{1}{2} x \cdot Cx - \frac{1}{2} (g \cdot x)^2 + \rho \, (g \cdot x) \text{ sat } \frac{\mu g \cdot x}{\rho} \tag{25}$$

(24) becomes

$$\Phi(x^o) = \int_0^\infty \left\{ -\dot{V} + (g \cdot x)^2 - \rho \, (g \cdot x) \text{ sat } \left(\frac{\mu g \cdot x}{\rho} \right) \right\} dt \tag{26}$$

$$= \frac{1}{2} x^o \cdot Bx^o - \int_0^\infty |g \cdot x| \left\{ \rho \left| \text{ sat } \left(\frac{\mu g \cdot x}{\rho} \right) \right| - |g \cdot x| \right\} dt$$

Thus the performance (in the sense of (24) is not degraded if

$$|g \cdot x| \left\{ \rho \left| \text{sat } \left(\frac{\mu g \cdot x}{\rho} \right) \right| - |g \cdot x| \right\} > 0 \tag{27}$$

Now, for

$$\left| \frac{\mu g \cdot x}{\rho} \right| \geq 1 \tag{28}$$

(27) is valid if

$$|g \cdot x| \leq \rho \tag{29}$$

which implies that $\mu \geq 1$. For

$$\left| \frac{\mu g \cdot x}{\rho} \right| \leq 1 \tag{30}$$

(27) is valid if

$$\mu \geq 1 \tag{31}$$

which implies that

$$|g \cdot x| \leq \mu |g \cdot x| \leq \rho \tag{32}$$

Summary of Results

Always sufficient for system stability	$\mu > 1/2;\ x \cdot Bx < \frac{-4\rho^2}{a \cdot g}$
Always sufficient for undegraded performance index	$\mu > 1;\quad \|g \cdot x\| \leq \rho$
Control is pure linear if	$\mu \|g \cdot x\| < \rho$
Control is bang-bang if	$\mu \|g \cdot x\| > \rho$

APPENDIX D

ON SYNTHESIS OF OPTIMAL BANG-BANG FEEDBACK CONTROL SYSTEMS WITH QUADRATIC PERFORMANCE INDEX

by

R.W. Bass and R.F. Webber
Space Systems Division
Hughes Aircraft Company
under

Contract No. NAS8-11421

Presented at 1965 JACC June 1965

Space Systems Division
AEROSPACE GROUP
Hughes Aircraft Company · Culver City, California

ON SYNTHESIS OF OPTIMAL BANG-BANG FEEDBACK CONTROL SYSTEMS WITH QUADRATIC PERFORMANCE INDEX

R. W. Bass and R. F. Webber, Guidance and Controls Division, Aerospace Group
Hughes Aircraft Company, Culver City, California

Abstract

This paper extends the work of Wonham and Johnson [1], who found the nature of the optimal control on the singular strip for a given quadratic performance index. Their solution required a special preliminary transformation to phase coordinates. In this work the optimal control is found, on the singular strip and in a neighborhood thereof, without the use of their special transformation. The optimal control law off the singular strip takes the form sgn $\sigma(x)$, where $\sigma(x)$ is a power series in the state vector, x. The terms of $\sigma(x)$ up to and including those of the third order are found.

Introduction

The work [1] is extended in several ways:

(i) Avoiding the change from state-variables to phase coordinates gives the present work complete generality, which is mistakenly claimed by [1]: the alleged reduction of the integrand of the quadratic performance index to a nonnegative definite weighted sum of squares in [1] is incorrect, in that some of the weighting coefficients may be negative, as simple examples show.

(ii) Computation of the plant's open-loop poles is avoided: the "singular control" gain vector is derived in terms of quadratic matricial equations closely related to those of optimal linear control;

(iii) The nature of the singular regime (linear) control in an (n-1)-dimensional strip near the origin) is completely explained by exhibiting a linearly switched bang-bang system, optimal near the origin, whose chattering regime [2] gives an average motion (the André-Seibert sliding regime*) which is identical with the singular regime;

(iv) A method for computing the coefficients of a multiple power-series in the state-variables which provides the local optimal control switching signal as an explicit feedback law is developed.

In [1], the optimal nonlinear control law is described implicitly by means of the familiar Hamiltonian Two-Point Boundary-Value Problem [4], whereas the present approach leads to an explicit solution of the equivalent Hamilton-Jacobi equation.

Principal Results

Let the system to be controlled have the state-vector form

$$\dot{x} = Ax + a\psi_o, \quad x(0) = x^o, \quad [A = (A_{ij}), \ a = (a_i)] \tag{1}$$

where the feedback control law $\psi_o = \psi_o(x)$ must satisfy

$$|\psi_o| \leq 1 \tag{2}$$

and, for some free terminal time T, $0 \leq T \leq +\infty$,

$$x(t) \to 0 \quad \text{as} \quad t \to T, \tag{3}$$

while at the same time minimizing the performance index

$$\Phi = \tfrac{1}{2}\int_0^T (x \cdot Cx)\,dt, \qquad (C = C^* > 0)^{\dagger} \tag{4}$$

where C is a given positive-definite symmetric matrix.

We shall prove that, in a sufficiently small neighborhood of x = 0, the optimal feedback control is precisely

$$\psi_o = \text{sgn}[\sigma], \qquad \sigma = g \cdot x, \tag{5}$$

(where as usual $\text{sgn}[\sigma] \equiv \sigma/|\sigma|$), where the gain vector g is defined as follows.

Let

$$\hat{P} = I_n - [1/(a \cdot Ca)]\, aa^*C, \tag{6}$$

so that $\hat{P}$ is a "projection matrix" with the properties

$$(\hat{P})^2 = \hat{P}, \qquad \hat{P}a = 0, \qquad (\hat{P})^*Ca = 0. \tag{7}$$

Let

$$B = B^* \geq 0, \qquad \det(B) = 0, \tag{8}$$

be a non-negative definite, singular symmetric matrix satisfying

$$B(A\hat{P}) + (A\hat{P})^*B - [1/(a \cdot Ca)]\, B(Aaa^*A^*)B = = -C + [1/(a \cdot Ca)]\, Caa^*C, \tag{9}$$

as well as the constraints

$$Ba = 0, \tag{10a}$$

$$(Aa) \cdot B(Aa) = -(Aa) \cdot Ca + \sqrt{a \cdot Ca} \tag{10b}$$

It was established in [1] that the condition of controllability [see (40) below] together with C positive definite are sufficient for (9) to have a solution B with the properties stated in (10).

Then the desired g will be given by

$$\boxed{g = -(1/\sqrt{a \cdot Ca})(BA + C)a,} \tag{11a}$$

which, by (8) and (10a), has as a corollary

$$\boxed{a \cdot g = -\sqrt{a \cdot Ca} < 0.} \tag{11b}$$

It should be noted that, by (11b), and [2], use of the control law (5), (11) must always lead to an end point x^1, namely, a time $T_* > 0$ and state x^1 such that

$$g \cdot x^1 = 0, \qquad |(A^*g) \cdot x^1| \leq -a \cdot g = |a \cdot g|, \tag{12a}$$

$$x(T_*) = x^1 \tag{12b}$$

As first noted by Flügge-Lotz [3], the solution x(t) of the system (1), (5), (11) cannot be defined

*Kliger [10] calls it the slippage regime.

†The notation of this paper will correspond as closely as possible to that of [7].

for $t > T_*$, because the relay $\psi_o = \mathrm{sgn}[\sigma]$ would begin to "chatter" at $\sigma = g \cdot x^1$. This difficulty was overcome in an elegant theory by André and Seibert [2], who assumed a small time-delay τ in the relay, namely $\psi(t) = \mathrm{sgn}[\sigma(t-\tau)]$, and derived the limit-motion $x(t)$, for $0 < T_* \le t < +\infty$, as $\tau \to 0$.

It turns out [2] that this <u>sliding motion</u> takes place in the hyperplane-strip

$$g \cdot x = 0, \qquad |(A^*g)\cdot x| \le -a \cdot g, \qquad (x = Px), \tag{13}$$

and is defined for $T_* \le t < +\infty$ by

$$\dot{x} = PAx, \qquad x(T_*) = x^1, \tag{14}$$

$$P = I_n - [1/(a \cdot g)]\, ag^*, \tag{15a}$$

$$Pa = 0, \qquad P^*g = 0. \tag{15b}$$

Note, however, by (15a), that

$$PA = A - [1/(a\cdot g)]ag^*A = A + a(-[1/(a\cdot g)]A^*g)^* \tag{16}$$

and that, by (13), $(|A^*g \cdot x| / |a \cdot g|) \le 1$.

Hence the sliding regime can be regarded (and, using <u>dual-mode</u> <u>control</u>, synthesized), as a <u>linear</u> control system of the form

$$\dot{x} = Ax + a\hat{\psi}_o, \quad g \cdot x = 0, \quad x(T_*) = x^1 \tag{17a}$$

$$\hat{\psi}_o = q \cdot x, \quad |\hat{\psi}_o| \le 1, \tag{17b}$$

$$q = -A^*g/(a \cdot g) + \kappa g \tag{17c}$$

$$-\infty < \kappa < \infty$$

We shall prove that the system (17) is <u>asymptotically</u> <u>stable</u> <u>on</u> $g \cdot x = 0$, and that

$$\Phi_* = \int_{T_*}^{+\infty} (x \cdot Cx)\,dt = x^1 \cdot Bx^1, \quad g \cdot x^1 = 0 \tag{18}$$

where B is given by (9) - (10), and where the feedback law $\hat{\psi}_o = q \cdot x$ actually minimizes Φ_* under the constraints $|\hat{\psi}_o| \le 1$, $g \cdot x^1 = 0$. Furthermore it will be proved that (17) is identical with the <u>singular</u> <u>regime</u> of Wonham and Johnson. [ADDED IN PROOF. Since this paper was accepted for presentation, the comments [10] of Kliger have appeared. Kliger makes a statement similar to point (iii) above, concerning implementation of the singular regime by means of a chattering regime. However, he does not mention the André-Seibert theory (14) - (15), and fails to prove the all-important results (11b) that $a \cdot g < 0$ and that the system (17) is asymptotically stable on $g \cdot x = 0$, which will follow from (22) below; the arguments of [10] do no more than prove that <u>if</u> $a \cdot g < 0$, the system (17) has <u>end-points</u> [2, 3] which is <u>necessary</u> for stability but <u>not</u> <u>sufficient</u>.]

Firstly, rewrite (11) as

$$\boxed{(a \cdot g)g = (BA + C)a, \quad (a \cdot g)^2 = a \cdot Ca.} \tag{19}$$

Now, using (19), and inserting (6) into (9), rearrange (9) to show its equivalence with

$$BA + A^*B = -C + gg^*. \tag{20}$$

By (10a) and (16),

$$BPA = BA. \tag{21}$$

Hence (20) can be expressed as

$$\boxed{B(PA) + (PA)^*B = -C + gg^*.} \tag{22}$$

Also, by (14) and (15b),

$$\frac{d}{dt}(g \cdot x) = g \cdot \dot{x} = g \cdot PAx = (P^*g)\cdot Ax \equiv 0, \tag{23}$$

so that $g \cdot x^1 = 0$ implies that

$$g \cdot x(t) \equiv 0, \quad (T_* \le t < +\infty). \tag{24}$$

Multiply (22) on the left by x^* and on the right by x, and note that, on $g \cdot x = 0$,

$$\Phi_* = x \cdot Bx. \tag{25}$$

By (14), (25) implies

$$\dot{\Phi}_* = x \cdot (BPA + A^*P^*B)x = -x \cdot Cx, \tag{26}$$

whence, integrating, one obtains

$$x(t) \cdot Bx(t) = x^1 \cdot Bx^1 - \int_{T_*}^{t} (x \cdot Cx)\,dt, \tag{27}$$

which shows that, on (24), $x(t) \to 0$ as $t \to +\infty$. This proves (18). The minimality of (18) was proved by Wonham and Johnson; hence, it remains only to identify (17) with their singular regime, which can be done by comparing (18) and (19) with the equations (23), (27), and (35) of [1].

By (19) and (17c),

$$q = -[1/(a\cdot g)]A^*g + \kappa g = -[1/(a\cdot g)]^2(A^*BA + A^*C)a + \kappa g \tag{28}$$

Hence, using (20) after multiplying on the right by A, (28) gives

$$\begin{aligned}(a \cdot Ca)q &= -(A^*BA + A^*C)a + \kappa g(a\cdot Ca) = \\ &= (BA^2 + CA - A^*C)a - (g \cdot Aa)g + \hat{\kappa} g\end{aligned} \tag{29}$$

$$\hat{\kappa} \doteq \kappa(a \cdot Ca)$$

or

$$\boxed{(a \cdot Ca)q = (BA^2 + CA - A^*C)a} \tag{30}$$

if

$$\boxed{g \cdot Aa = -1.} \quad \kappa = -1/a\cdot Ca \tag{31}$$

But (31) is a consequence of (10b) and (19). Hence, (29) holds. The equations (19) and (30) give g and q according to [1], while the preceding argument has shown their consistency with (17c).

It is easy to give a direct proof that (17), (19), (22), (30), (31) correspond to singular control. By (1) and (4), and the Maximum Principle [4], define a Hamiltonian by

$$\mathcal{H} = x \cdot A^*y + (a \cdot y)\hat{\psi}_o - \tfrac{1}{2} x \cdot Cx. \tag{32}$$

The Hamiltonian canonical (variational) equations associated with (32) are

$$\dot{x} = Ax + a\hat{\psi}_o = \mathrm{grad}_{(y)}\mathcal{H}, \tag{33a}$$

$$\dot{y} = -A^*y + Cx = -\mathrm{grad}_{(x)}\mathcal{H}, \quad y(T_*) = y^1. \tag{33b}$$

Singular control occurs when $\mathcal{H} \equiv 0$ by virtue of, separately,

$$a \cdot y \equiv 0, \tag{34a}$$

$$x \cdot A^*y \equiv \tfrac{1}{2} x \cdot Cx. \tag{34b}$$

Now assume that

$$y = -Bx, \quad (a \cdot y^1 = 0, \quad Ba = 0). \tag{35}$$

Clearly $a \cdot y = -a \cdot Bx = -(Ba)\cdot x \equiv 0$.

Similarly, by (20) and (35), equation (34b) holds on the strip (13). But we have already proved that

$\hat{\psi}_o = q \cdot x$ implies that the system (33a) is asymptotically stable on (13). Furthermore, by (31), (19), and (30), equations (33a, b) and (35) imply that

$$a \cdot \dot{y} \equiv (a \cdot g)(g \cdot x) , \qquad (36a)$$

$$a \cdot \ddot{y} \equiv [(a \cdot Ca)q - (BA^2 + CA - A^*C)a] \cdot x \equiv 0 . \qquad (36b)$$

Hence $g \cdot x^1 = 0$ implies that $a \cdot \dot{y} \equiv 0$, whence (35) is in fact a solution of (33a, b)-(34a, b) for $\hat{\psi}_o = q \cdot x$.

Optimal trajectories outside of the singular strip (13) can be obtained by the flooding technique, as noted in [1]. Since application of the results of [1] require a special coordinate system, whereas the present formulation is completely general, the flooding procedure will be described in the present notation. (A discretized version of flooding is well known as Dynamic Programming [5].) Since $a \cdot y$ is to vanish only at isolated times, the Maximum Pr iple [4] is applicable. The optimal trajectories leading to the singular strip are generated by starting on the strip and integrating the Hamiltonian (canonical) equations backwards in time. If the terminal state x^1 is an arbitrary point on the singular strip, then the corresponding terminal co-state $y^1 = -Bx^1$ is known by the preceding characterization of the matrix B. The (optimized) Hamiltonian is therefore

$$\mathcal{H} = x \cdot A^* y + |a \cdot y| - \tfrac{1}{2} x \cdot Cx , \qquad (37a)$$

where the result

$$\psi_o = \mathrm{sgn}[a \cdot y] \qquad (37b)$$

is a consequence of $a \cdot y \neq 0$; now integrate

$$-\dot{x} = Ax + a\,\mathrm{sgn}[a \cdot y] = \mathrm{grad}_{(y)}\mathcal{H}, \quad x(0) = x^1, \qquad (37c)$$

$$-\dot{y} = -A^* y + Cx = -\mathrm{grad}_{(x)}\mathcal{H} , \quad y(0) = y^1 = -Bx^1, \qquad (37d)$$

for $0 \le t < +\infty$. Every state $x(t)$ attained in this manner will have as its co-state the associated $y(t)$, and the optimal control value $\psi_o = \mathrm{sgn}[a \cdot y(t)]$.

The preceding flooding method (37) is just a technique for solving the Hamilton-Jacobi partial differential equation, $\mathcal{H} = 0$, by the method of characteristics. In fact, outside of the singular strip one has

$$\psi_o = \mathrm{sgn}[a \cdot y] = -\mathrm{sgn}[(a \cdot \mathrm{grad}\, \Phi)] \qquad (38a)$$

$$y = -\mathrm{grad}_{(x)}\Phi , \quad \Phi = \Phi(x) , \qquad (38b)$$

$$\mathcal{H} = -Ax \cdot \mathrm{grad}\,\Phi + |a \cdot \mathrm{grad}\,\Phi| - \tfrac{1}{2} x \cdot Cx \equiv 0 . \qquad (38c)$$

These equations can be re-written in the form

$$(Ax + a\epsilon_o) \cdot \mathrm{grad}\,\Phi = -\tfrac{1}{2} x \cdot Cx , \qquad (39a)$$

$$\epsilon_o = -\mathrm{sgn}[a \cdot \mathrm{grad}\,\Phi] , \quad (\epsilon_o^2 = 1) , \qquad (39b)$$

$$\mathrm{grad}\,\Phi = Bx \quad \text{when} \quad g \cdot x = 0 , \quad |A^* g \cdot x| \le |a \cdot g| \qquad (39c)$$

where, in the notation of stability theory, (39a) is equivalent to

$$\dot{\Phi} = -\tfrac{1}{2} x \cdot Cx , \qquad (39d)$$

namely, Φ is a positive-definite Liapunov function whose Lie derivative is the negative-definite function $-(1/2) x \cdot Cx$, and where the integration of the partial differential equations (39a) - (39b) is to be performed subject to the boundary conditions (39c).

Numerous publications in this field have stated that explicit solution of (39) is a "hopeless" task. However, by using some new results of Bass [6] (see also [7] - [9]), the partial differential equation (39) can be solved explicitly in a neighborhood of $x = 0$, as will be shown.

The computation of certain auxiliary vectors and matrices is a preliminary step.

Define the controllability matrix (Kalman) by

$$D = (a, \ Aa, \ A^2 a, \cdots, A^{n-1} a) \qquad (40)$$

and assume, as was done implicitly in assuming solubility of (9) - (10), that $\det D \neq 0$. Then a vector b exists which is defined by

$$b = (D^{-1})^* e^n , \quad (D^* b = e^n = (0, 0, \cdots, 0, 1)^*) . \qquad (41)$$

By definition, b has the property that $b \cdot A^{i-1} a \equiv [(A^*)^{i-1} b] \cdot a = \delta_{in}$. ($\delta_{nn} \triangleq 1$; $\delta_{in} \triangleq 0$, $i \neq n$.)

It can be shown [7] that if one defines a phase variable θ_1 by setting

$$\theta_1 = b \cdot x , \quad \left(\theta_1^{[i]} = d^i \theta_1 / dt^i = [(A^*)^i b] \cdot x , \quad (i = 0, 1, \cdots, n-1)\right) \qquad (42)$$

then the system (1) is equivalent to

$$\Delta(d/dt)\theta_1 \equiv \sum_{i=0}^{n} \alpha_i \theta_1^{[i]} = \psi_o , \qquad (43)$$

(which is in terms of the phase-coordinates $\theta_1, \dot{\theta}_1, \cdots, \theta_1^{[n-1]}$, instead of the state variables $x_1, \cdots, x_n$) where the α_i are defined by

$$\Delta(s) = \det(sI_n - A) \equiv \sum_{i=0}^{n} \alpha_i s^i , \qquad (44)$$

and where the inverse of the change of variables (42) is given explicitly [7] by

$$x = \{\Gamma(d/dt)\theta_1\} a \equiv \sum_{i=1}^{n} \theta_1^{[i-1]} S_i a , \qquad (45)$$

$$\Gamma(s) \equiv \sum_{i=1}^{n} s^{i-1} S_i . \qquad (46)$$

Here, if $\mathcal{L}$ denotes Laplace transform, and s the complex Laplace variable, the polynomial $\Delta(s)$ is the open-loop characteristic polynomial of (1), and the matrix $\Gamma(s)$ is the numerator of the open-loop matrix transfer function $G(s)$, which is given by

$$G(s) \equiv (sI_n - A)^{-1} = \mathcal{L}\{e^{At}\} = \frac{\Gamma(s)}{\Delta(s)} , \qquad (47)$$

where the theoretical definition of the matrices S_i is

$$S_i = \sum_{j=i}^{n} \alpha_j A^{j-i} , \quad (i = 1, 2, \cdots, n) . \qquad (48)$$

The definition (44) requires n! multiplications and for large n cannot be used to compute α_i. However, the coefficients α_i and matrices S_i can be computed in about n^4 multiplications* by the algorithm†:

*For example, a Hughes computer subroutine finds (α_i, S_i) for a 10×10 matrix in about 5 seconds of IBM 7094 time. † Leverrier (1840); cf. [7].

$$\alpha_n = 1 \ , \quad S_n = I_n \ , \tag{49a}$$

$$\alpha_{n-j} = -(1/j)\operatorname{trace}(AS_{n-j+1}) \ , \quad S_{n-j} = \alpha_{n-j}I_n + AS_{n-j+1} \ ,$$
$$(j = 1, 2, \cdots, n) \tag{49b}$$

whose accuracy can be checked by the fact that $S_o \equiv 0$ should hold (Cayley-Hamilton theorem).

Now define vector transfer functions v(s), u(s) by

$$v(s) = \Gamma^*(s)b = \sum_{i=1}^{n}\sum_{k=i}^{n} s^{i-1}\alpha_k (A^*)^{k-i}b =$$

$$= \sum_{j=1}^{n}\left(\sum_{k=j}^{n}\alpha_k s^{k-j}\right)(A^*)^{j-1}b \ , \tag{50}$$

$$u(s) = \Gamma(s)a/\Delta(s) = \sum_{i=1}^{n}\sum_{j=i}^{n}\left[\frac{s^{i-1}}{\Delta(s)}\right]\alpha_j A^{j-i}a =$$

$$= \sum_{j=1}^{n}\left(\sum_{k=j}^{n}\alpha_k\left[\frac{s^{k-j}}{\Delta(s)}\right]\right)A^{j-1}a \ , \tag{51}$$

and note that (by the Cayley-Hamilton theorem which gives $(sI_n - A)\Gamma(s) \equiv \Delta(s)I_n$), the following results are identities:

$$A^*v(s) \equiv sv(s) - \Delta(s)b \ , \quad Au(s) \equiv su(s) - a \ , \tag{52a}$$

$$v(s)\cdot a \equiv 1 \ , \quad u(s)\cdot b \equiv 1/\Delta(s) \ . \tag{52b}$$

Also, it can be proved [7] that

$$D^{-1} \equiv (a, Aa, \cdots, A^{n-1}a)^{-1} = (S_1^*b, \cdots, S_n^*b)^* \ , \tag{53}$$

and, analogously,

$$(b, A^*b, \cdots, (A^*)^{n-1}b)^{-1} = (S_1a, \cdots, S_na)^* \ . \tag{54}$$

Hence, if the vector transfer function w(s) is defined by

$$w(s) = (S_1a, \cdots, S_na)^*v(s) \ , \tag{55}$$

it will be true that [multiply (50) by (54)]

$$w(s) = \sum_{j=1}^{n}\left(\sum_{k=j}^{n}\alpha_k s^{k-j}\right)e^j \ , \tag{56}$$

where the e^j are the fundamental unit vectors $(I_n = (e^1, \cdots, e^n))$.

Next, compute a sequence of numbers $\{\beta_i\}$ recursively by

$$\beta_{-\nu} = 0 \ , \quad (\nu = 1, 2, \cdots, n-1) \ ; \quad \beta_o = 1 \ , \tag{57a}$$

$$\beta_\nu = -\sum_{i=0}^{\nu-1}\alpha_{i+n-\nu}\beta_i \ , \quad (\nu = 1, 2, \cdots, n) \ , \tag{57b}$$

$$\beta_{n+\nu} = -\sum_{i=\nu}^{n+\nu-1}\alpha_{i-\nu}\beta_i \ , \quad (\nu = 1, 2, 3, \cdots) \ , \tag{57c}$$

and note the resultant identities [7], [8]

$$1/\Delta(s) \equiv \sum_{j=0}^{\infty}\beta_j s^{-(n+j)}, \ (\text{for } |s| > \rho_o \text{ defined by (66)}) \ , \tag{58}$$

$$b\cdot A^{j-1}a = \beta_{j-n} \ , \quad (j = 1, 2, 3, \cdots) \ . \tag{59}$$

Define a new set of state variables $\phi_i = \phi_i(x)$, $(i = 1, 2, \cdots, n)$, by

$$D^{-1}x = (\phi_1, \phi_2, \cdots, \phi_n)^* \ , \tag{60a}$$

and note that by (53) and (42),

$$x = \sum_{i=1}^{n}\phi_i A^{i-1}a \ , \quad \phi_i = (S_i^*b)\cdot x \ , \quad \theta_1 = \phi_n \ . \tag{60b}$$

Further note that by (46), (50), (53) and (60a),

$$\xi_o(s) \triangleq v(s)\cdot x = \sum_{i=1}^{n}\phi_i s^{i-1} \ . \tag{61}$$

Since s is an arbitrary complex variable, (61) is equivalent to the set of conditions obtained by equating the coefficients of like powers of s, i.e., (61) is just a condensed statement of (60b).

Now multiply (1) scalarly by v(s) and use (52a) and (42) in order to verify that if x = x(t) satisfies (1), then $\xi_o = \xi_o(s, t) = v(s)\cdot x(t)$ satisfies

$$d\xi_o(s)/dt = s\xi_o(s) + \psi_o - \Delta(s)\phi_n \ . \tag{62a}$$

Again, (62) is just a condensed statement of the differential equations

$$\dot{\phi}_1 = -\alpha_o\phi_n + \psi_o \ , \tag{62b}$$

$$\dot{\phi}_i = \phi_{i-1} - \alpha_{i-1}\phi_n \ , \quad (i = 2, 3, \cdots, n) \ , \tag{62c}$$

obtained by equating coefficients of like powers of s.

The motivation for the preceding derivation of (62a) is that, in the special case when the roots $\lambda_1, \cdots, \lambda_n$ of $\Delta(s) = 0$ are distinct, one can define

$$v^i = v(\lambda_i) \ , \quad u^i = \{[\Delta(s)/\Delta'(s)]u(s)\}_{s=\lambda_i} \ ,$$
$$(i = 1, 2, \cdots, n) \ , \tag{62d}$$

$$\xi_i = v^i\cdot x = \xi(\lambda_i) \ , \tag{62e}$$

and obtain from (52),

$$A^*v^i = \lambda_i v^i \ , \quad v^i\cdot a = 1 \ ; \quad Au^i = \lambda_i u^i \ ,$$
$$u^i\cdot b = 1/\Delta'(\lambda_i) \ , \quad (i = 1, 2, \cdots, n) \tag{62f}$$

while (62a) becomes the Lur'e canonical form

$$\dot{\xi}_i = \lambda_i\xi_i + \psi_o \ , \quad (i = 1, 2, \cdots, n) \ . \tag{62g}$$

Thus, (62a) is a generalization of the Lur'e canonical form, valid whether the roots λ_i of $\Delta(s)$ are distinct or not.

Next, assume that ψ_o is piecewise constant, and, in particular that

$$\psi_o = \epsilon_o \ , \quad \epsilon_o^2 = 1 \ , \quad (\epsilon_o = +1 \text{ or } \epsilon_o = -1) \ . \tag{63}$$

If ξ_o satisfies (62), (63), it will henceforth be called $\xi_o = \xi_o(s;\epsilon_o)$. Define a function

$$\kappa = \kappa(s, \xi_o) \equiv (\epsilon_o/s)\log[1 + \epsilon_o s\xi_o(s;\epsilon_o)] \ , \tag{64}$$

and note that κ is analytic in x_i for

$$\|x\| < 1/\rho_o \bar{v}(\rho_o) , \quad |s| \le \rho_o , \tag{65a}$$

$$\bar{v}(\rho_o) \equiv \sum_{j=1}^{n}\left(\sum_{k=j}^{n} |\alpha_k| \rho_o^{k-j}\right)\|A\|^{j-1}\|b\| . \tag{65b}$$

For future reference, choose

$$\rho_o > \|A\| \quad (\ge \max_{i=1,\cdots,n} |\lambda_i| , \quad \Delta(\lambda_i) = 0) . \tag{66}$$

Referring to the n new state variables $\phi_i = (S_i^* b) \cdot x$ of (60) and (62b, c), note that they can be related to the state variables x_i and phase coordinates $\theta_1^{[i-1]}$, $(i = 1, 2, \cdots, n)$ by the following explicit and uniquely invertible transformations [7],

$$\phi_i = \sum_{j=1}^{n} \alpha_j \theta_1^{[j-i]} , \quad (i = 1, 2, \cdots, n) , \tag{67a}$$

$$x = \sum_{i=1}^{n} \phi_i A^{i-1} a = \sum_{i=1}^{n} \theta_1^{[i-1]} S_i a , \tag{67b}$$

$$\theta_1^{[i-1]} = \sum_{j=0}^{i-1} \beta_j \phi_{j+n-i+1} , \quad (i = 1, 2, \cdots, n) , \tag{67c}$$

and that the function $\kappa(s, \xi_o)$ can be [8] expressed in terms of the $\phi_i = \phi_i(x)$ as follows. Define $\omega_i = \omega_i(x, \epsilon_o)$ recursively by

$$\omega_1 = \phi_1 , \tag{68a}$$

$$\omega_\nu = \phi_\nu - \frac{\epsilon_o}{\nu} \sum_{j=1}^{\nu-1} (\nu - j)\omega_{\nu-j}\phi_j =$$

$$= \phi_\nu - \frac{\epsilon_o}{\nu} \sum_{j=1}^{\nu-1} j\omega_j \phi_{\nu-j} , \quad (\nu = 2, 3, \cdots, n) \tag{68b}$$

$$\omega_{n+\nu} = - \frac{\epsilon_o}{n+\nu} \sum_{j=1}^{n} (n + \nu - j)\omega_{n+\nu-j}\phi_j =$$

$$= - \frac{\epsilon_o}{n+\nu} \sum_{j=\nu}^{n+\nu-1} j\omega_j \phi_{n+\nu-j} , \quad (\nu = 1, 2, 3, \cdots) . \tag{68c}$$

Then it can be proved [6], [7], [8] that

$$\kappa(s, \xi_o(s;\epsilon_o)) \equiv (\epsilon_o/s)\log[1 + \epsilon_o s \xi_o(s;\epsilon_o)] =$$

$$= \sum_{\nu=1}^{\infty} \omega_\nu(x, \epsilon_o)s^{\nu-1} . \tag{68d}$$

Now we are in a position to define certain very important functions $\sigma_j = \sigma_j(x;\epsilon_o)$ by

$$\sigma_j = \sigma_j(x;\epsilon_o) = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho_o} \left[\frac{s^{j-1}}{\Delta(s)}\right] \kappa(s, \xi_o(s;\epsilon_o))\, ds ,$$

$$(j = 1, 2, \cdots, n) , \tag{69}$$

and note that the $\sigma_j(x;\epsilon)$ are analytic in x on (65). Using (58) and (68), it is easy to prove by the calculus of residues that

$$\sigma_j(x;\epsilon_o) = \sum_{i=0}^{\infty} \beta_i \omega_{i+n-j+1}(x, \epsilon_o) =$$

$$= \theta_1^{[j-1]} + O(\|x\|^2) =$$

$$= [(A^*)^{j-1} b] \cdot x - \tfrac{1}{2}\epsilon_o (x \cdot Q_j x) + O(\|x\|^3) , \tag{70a}$$

$$Q_j = (A^*)^{j-1} Q_1 , \quad Q_1 = (D^{-1})^* E D^{-1} , \tag{70b}$$

where the i, $j^{\underline{th}}$ element of E is defined by

$$e^i \cdot E e^j = \beta_{i+j-n} , \quad (i, j = 1, 2, \cdots, n) . \tag{70c}$$

When the λ_i are distinct,

$$b = \sum_{i=1}^{n} \left[\frac{1}{\Delta'(\lambda_i)}\right] v^i \tag{70d}$$

$$Q_1 = \sum_{i=1}^{n} \left[\frac{\lambda_i}{\Delta'(\lambda_i)}\right] v^i (v^i)^* . \tag{70e}$$

Now define the nonlinear vector function

$$p = p(x;\epsilon_o)$$

by

$$p = p(x;\epsilon_o) = (\sigma_1(x;\epsilon_o), \cdots, \sigma_n(x;\epsilon_o))^* , \tag{71}$$

and note that

$$p = p(x;\epsilon_o) = (b, A^* b, \cdots, (A^*)^{n-1} b)^* x + O(\|x\|^2) . \tag{72}$$

Consequently, the transformation

$$\sigma = p(x;\epsilon_o) , \quad \sigma \equiv (\sigma_1, \sigma_2, \cdots, \sigma_n)^* , \tag{73}$$

has a unique inverse

$$x = h(\sigma;\epsilon_o) = (S_1 a, \cdots, S_n a)\sigma + O(\|\sigma\|^2) , \tag{74}$$

for all $\|x\|$ sufficiently small. It can be shown [8] that this inverse is given explicitly by

$$x = h(\sigma;\epsilon_o) = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho_o} \frac{\epsilon_o}{s}\{\exp[\epsilon_o w(s) \cdot \sigma] - 1\} u(s)\, ds =$$

$$= \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho_o} \frac{\epsilon_o}{s}\left\{\exp\left[\epsilon_o \sum_{j=1}^{n}\left(\sum_{k=j}^{n} \alpha_k s^{k-j}\right)\sigma_j\right] - 1\right\}\left(\sum_{i=1}^{n}\left[\frac{s^{i-1}}{\Delta(s)}\right] S_i a\right) ds , \tag{75}$$

and that the transformations (73) - (74) are reciprocal for all x on (65), i.e., that

$$x \equiv h(p(x;\epsilon_o);\epsilon_o) , \quad \sigma \equiv p(h(\sigma;\epsilon_o);\epsilon_o) \tag{76}$$

for all $\|x\| < 1/\rho_o \bar{v}(\rho_o)$. Furthermore, it can be proved [7] that

$$\det(h_\sigma(0;\epsilon)) = \det(D) , \tag{77}$$

which re-emphasizes the fact that the condition of controllability $\det(D) \neq 0$ plays an essential role in the construction of (69) - (71), and (75) as reciprocal transformations; in other words, controllability is sufficient for the functions $\sigma_j = \sigma_j(x;\epsilon_o)$ of (69) - (70) to be "functionally independent".

The significance of the σ_j's is that $\sigma_1, \sigma_2, \cdots, \sigma_{n-1}$ are <u>first integrals</u> of

$$\dot{x} = Ax + a\epsilon_o \ , \quad (\epsilon_o^2 = 1) \tag{78}$$

while σ_n is an <u>isochrone</u> [8]. That is, by substituting (78) - or its equivalent, (62) - into (64) and (69), it is easy to verify that

$$\dot{\sigma}_j(x;\epsilon_o) \equiv d\sigma_j(x;\epsilon_o)/dt = \epsilon_o \delta_{jn} = \begin{cases} 0, & j = 1, 2, \cdots, n-1 \\ 1, & j = n \ . \end{cases} \tag{79}$$

Another interpretation of (79) is the equivalence

$$\boxed{\dot{x} = Ax + a\epsilon_o \ , \quad x = h(\sigma) \ , \Longleftrightarrow \dot{\sigma} = \epsilon_o e^n \ , \quad \sigma = p(x).} \tag{80}$$

In other words, the <u>nonlinear</u> change of state variables

$$x_i = h_i(\sigma_1, \cdots, \sigma_n; \epsilon_o) \ , \quad \sigma_i = p_i(x;\epsilon_o) \ , \quad (i = 1, 2, \cdots, n) \ , \tag{81}$$

"rectifies" the phase-portrait of (78) for $\|x\| \leq 1/\rho_o \, \bar{v}(\rho_o)$ by transforming the "streamlines" of (78) into parallel straight lines.

Using (80), it is possible to solve the Hamilton-Jacobi equation

$$\boxed{\dot{\Phi} \equiv (Ax + a\epsilon_o)\cdot \mathrm{grad}_{(x)}\Phi = -\Psi(x) \ , \quad \Phi(0) = 0 \ ,} \tag{82}$$

by noting its equivalence to

$$\dot{\Phi} \equiv \epsilon_o e^n \cdot \mathrm{grad}_{(\sigma)}\Phi = -\Psi(h(\sigma)) \ , \quad \Phi(0) = 0 \ , \tag{83a}$$

or

$$\boxed{\partial\Phi(\sigma_1, \cdots, \sigma_n)/\partial\sigma_n = -\epsilon_o \Psi(h(\sigma_1, \cdots, \sigma_n)) \ .} \tag{83b}$$

In the present case, $\Psi = (1/2)x \cdot Cx$. Hence the <u>general solution of (83) is</u> [clearly $\partial\Phi_o/\partial\sigma_n \equiv 0$]

$$\Phi = \Phi_o + \Phi_1 \ , \tag{84a}$$

$$\Phi_o = \Phi_o(\sigma_1, \sigma_2, \cdots, \sigma_{n-1}) \ , \quad \Phi_1 = \Phi_1(\sigma_1, \cdots, \sigma_n) \tag{84b}$$

$$\Phi_1 \equiv -(1/2)\epsilon_o \int_0^{\sigma_n} [h(\sigma_1, \cdots, \sigma_{n-1}, z_n) \cdot Ch(\sigma_1, \cdots, \sigma_{n-1}, z_n)] dz_n , \tag{84c}$$

<u>where Φ_o is an arbitrary function.</u>

It remains only to specify $\Phi_o = \Phi_o(\sigma_1, \cdots, \sigma_{n-1})$ in order to have found the solution of (82) as

$$\Phi = \Phi_o(\sigma_1(x;\epsilon_o), \cdots, \sigma_{n-1}(x;\epsilon_o)) + \Phi_1(\sigma_1(x;\epsilon_o), \cdots, \sigma_n(x;\epsilon_o)) \ . \tag{85}$$

However, the preceding theory of the singular strip has shown [(1) - (37)] that if

$$\psi_o = \psi_o(x) = \epsilon_o \quad \text{for} \quad 0 \leq t \leq T_* \ , \quad \text{and}$$

$$g \cdot x(T_*) = 0 \ , \quad |A^*g \cdot x(T_*)| \leq |a \cdot g| \ ,$$

and

$\psi_o = q \cdot x \ , \quad q = -(a \cdot g)^{-1} A^*g$ +xg for $T_* < t < +\infty$,

then [integrating (82) for $0 \leq t \leq T_*$ and adding (18)]

$$\Phi = \Phi(x^o) = \tfrac{1}{2}\int_0^{+\infty} [x(t) \cdot Cx(t)] \, dt \tag{86a}$$

is an expression for the solution of

$$\dot{\Phi} \equiv (Ax + a\psi_o) \cdot \mathrm{grad}_{(x)}\Phi = -(1/2)x \cdot Cx \tag{86b}$$

in a neighborhood of the points $\{x(t) \,|\, 0 \leq t < +\infty\}$.

Hence (by the "Principle of Optimality" [5])

$$\Phi(x^1) = \Phi(x(T_*)) = (1/2)x^1 \cdot Bx^1 \tag{87}$$

and so (85) must have the boundary values

$$\Phi(x) = (1/2)x \cdot Bx, \quad \text{when } g \cdot x = 0 \ , \quad |A^*g \cdot x| \leq |a \cdot g| \ . \tag{88}$$

By (74) and $S_n = I_n$,

$$x = h(\sigma_1, \cdots, \sigma_n) = (S_1 a)\sigma_1 + \cdots + (S_{n-1}a)\sigma_{n-1} + a\sigma_n + O(\|p\|^2) \ , \tag{89}$$

whence, by (11b)

$$g \cdot x = (g \cdot S_1 a)\sigma_1 + \cdots + (g \cdot S_{n-1}a)\sigma_{n-1} - \left(\sqrt{a \cdot Ca}\right)\sigma_n + O(\|p\|^2) \ . \tag{90}$$

Since $a \cdot Ca \neq 0$, the standard expression for the reversion of power series applies to give an analytic function $\Phi_n = \Phi_n(\sigma_1, \cdots, \sigma_{n-1})$ such that

$$g \cdot x = 0 \Longleftrightarrow \sigma_n = \Phi_n(\sigma_1, \cdots, \sigma_{n-1}) \ , \tag{91a}$$

$$\sigma_n = [(g \cdot S_1 a)/\sqrt{a \cdot Ca}]\sigma_1 + \cdots + [(g \cdot S_{n-1}a)/\sqrt{a \cdot Ca}]\sigma_{n-1} + O(\|\sigma - \sigma_n e^n\|^2) \ . \tag{91b}$$

Hence if $\|x\| \leq \min(1/\rho_o \bar{v}(\rho_o), |a \cdot g|/\|A^*g\|)$, the desired function Φ_o is [by (88) and (91)] given explicitly by

$$\Phi_o = (1/2)h(\sigma_1, \cdots, \sigma_{n-1}, \Phi_n) \cdot Bh(\sigma_1, \cdots, \sigma_{n-1}, \Phi_n) + - \Phi_1(\sigma_1, \cdots, \sigma_{n-1}, \Phi_n) \ . \tag{92}$$

Using (84) and (92), we may define an analytic function

$$\sigma_o = \sigma_o(x;\epsilon_o) = a \cdot \mathrm{grad}_{(x)}\Phi \ , \quad \Phi = \Phi_o + \Phi_1 \ , \tag{93}$$

such that, by the Maximum Principle, in the regions $\epsilon_o \neq 0$ it will be true that

$$\epsilon_o = -\mathrm{sgn}[\sigma_o(x;\epsilon_o)] \ , \quad (\epsilon_o^2 = 1) \ . \tag{94}$$

Note that σ_o is a multiple power-series jointly in the $(n+1)$ variables ϵ_o and x_i; because $\epsilon_o^2 = 1$, the series for σ_o can be collected into the sum of two series having the form

$$\sigma_o = \sigma_o(x;\epsilon_o) = \hat{\sigma}_o(x) + \epsilon_o \tilde{\sigma}_o(x) \ , \tag{95}$$

where $\hat{\sigma}_o(x)$ and $\tilde{\sigma}_o(x)$ are analytic functions of x, independent of ϵ_o. If the inequality

$$-\infty < |\tilde{\sigma}_o(x)| \leq |\hat{\sigma}_o(x)| \tag{96}$$

defines a region which contains a neighborhood of $x = 0$ (i.e., $\|x\| < \gamma_o$, for some $\gamma_o > 0$), then in this neighborhood (with its intersection with the singular strip deleted) the optimal control law will be given by

$$\epsilon_o = - \operatorname{sgn}[\hat{\sigma}_o(x)] . \tag{97}$$

An alternative approach to finding the switching function makes use of the fact that, outside of the singular strip, the switching surface must be a first integral of (78), for an appropriate $\epsilon_o = \pm 1$. Hence the surface must be given by branches of

$$\Phi(x;\epsilon_o) \equiv \hat{\Phi}(\sigma_1(x;\epsilon_o), \cdots, \sigma_{n-1}(x;\epsilon_o)) = 0 , \tag{98}$$

where $\hat{\Phi} = \hat{\Phi}(z_1, \cdots, z_{n-1})$ is a suitable analytic function. (Just use (82), with $\Psi \equiv 0$, to get $\Phi = 0$.) The function $\hat{\Phi}$ can be determined from the obvious boundary condition that the set of σ_i such that

$$\hat{\Phi}(\sigma_1, \cdots, \sigma_{n-1}) = 0 , \qquad \sigma_n = \Phi_n(\sigma_1, \cdots, \sigma_n) \tag{99a}$$

must contain the set ("edges" of the singular strip)

$$\sigma_n = \Phi_n(\sigma_1, \cdots, \sigma_n), \quad (A^*g) \cdot h(\sigma_1, \cdots, \sigma_{n-1}, \sigma_n) = \pm (a \cdot g) . \tag{99b}$$

Hence we can choose

$$\hat{\Phi}_\pm \equiv (A^*g) \cdot h(\sigma_1, \cdots, \sigma_{n-1}, \Phi_n(\sigma_1, \cdots, \sigma_{n-1})) \pm (a \cdot g) . \tag{100}$$

The use of (93) - (97) will now be illustrated by an explicit power series expansion through terms cubic in the x_i.

Using (74), (91b), and (70a), we have from (93) and (54)

$$\begin{aligned} \sigma_o &= e^n \cdot \operatorname{grad}_{(\sigma)}\Phi = \partial\Phi_1/\partial\sigma_n = -\epsilon_o \Psi(h(\sigma_1, \cdots \sigma_n)) = \\ &= -(1/2)\epsilon_o \sum_{i=1}^{n} \sum_{j=1}^{n} [a \cdot S_i^* C S_j a]\sigma_i \sigma_j + O(\|\sigma\|^3) = \\ &= \hat{\sigma}_o + \epsilon_o \tilde{\sigma}_o \end{aligned} \tag{101}$$

where [after some algebraic manipulations]

$$\tilde{\sigma}_o = -(1/2) x \cdot Cx + O(\|x\|^3) , \tag{102}$$

$$\hat{\sigma}_o = - \sum_{i=1}^{n} (x \cdot Q_i x)([CS_i a] \cdot x) + O(\|x\|^4) . \tag{103}$$

Assuming now that the lowest order terms determine the relative signs of $\tilde{\sigma}_o$ and $\hat{\sigma}_o$, it is clear from (93) - (97) that

$$\epsilon_o = \operatorname{sgn}\left[\sum_{i=1}^{n} (x \cdot Q_i x)([CS_i a] \cdot x) \right] , \tag{104}$$

whenever

$$x \cdot Cx < 2 \left| \sum_{i=1}^{n} (x \cdot Q_i x)([CS_i a] \cdot x) \right| . \tag{105}$$

Conclusions and Epilogue

The properties of the optimal control in a neighborhood have been described. This control is obtained by using certain closed form nonlinear transformations. The method is an analytical version of obtaining optimal trajectories off the singular strip by the method of flooding. It should be emphasized that the solution to the problem presented is local in nature and could be investigated by simulation to find out empirically its global validity.

The authors have recently discovered how to find the optimal control for stable plants off the singular strip for performance indices of the form

$$\int_0^T \psi_{2\nu}(x)\, dt ,$$

where $\psi_{2\nu}(x) \geq 0$, $\nu = 2, \cdots$, and $\psi_{2\nu}(\mu x) = \mu^{2\nu} \psi_{2\nu}(x)$.

That is, $\psi_{2\nu}$'s are positive semi-definite homogeneous multinomial forms of degree 2ν. It is interesting to note that the optimal control off the strip in this case is of the form

$$- \operatorname{sgn}[a \cdot \operatorname{grad}_{(x)} \phi_{2\nu}(x)]$$

where $\phi_{2\nu}(x)$ is a positive semi-definite homogeneous multinomial form. That is, the argument of sgn consists of only one term of a power series and not an entire power series. These topics will be discussed in a forthcoming article.

References

1. W.M. Wonham and C.D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," Journal of Basic Engineering, Series D, March 1964, pp. 107-115.
2. J. André and P. Seibert, "Piecewise Continuous Differential Equations," Boletin de la Sociedad Matematica Mexicana, 1961, pp. 242-245.
3. I. Flügge-Lotz, Discontinuous Automatic Control, Princeton, 1953.
4. L.S. Pontryagin, et. al., The Mathematical Theory of Optimal Processes, Interscience, 1962.
5. R. Bellman et. al., Some Aspects of the Mathematical Theory of Control Processes, The Rand Corp. Project Rand R-313 (January 16, 1958).
6. R.W. Bass - Final Reports of Contracts: AFOSR 2754 (1962); NASr-25 (1961); NAS8-1628 (1962); Lecture Notes, Langley Research Center, NASA, (1961).
7. R.W. Bass and I. Gura, "Linear Canonical Forms for Controllable Systems," (to appear).
8. R.W. Bass and I. Gura, "A Nonlinear Canonical Form for Controllable Systems," (to appear).
9. R.W. Bass and I. Gura, "Synthesis of Time-Optimal Feedback Control Systems," (to appear).
10. I. Kliger, "Discussion on the Stability of the Singular Trajectory with Respect to 'Bang-Bang' Control," IEEE Transactions on Automatic Control, Vol. AC-9 (1964), pp. 582-585.

APPENDIX E

OPTIMAL NONLINEAR FEEDBACK CONTROL DERIVED FROM QUARTIC AND HIGHER-ORDER PERFORMANCE CRITERIA

by

R. W. Bass and R. F. Webber
Space Systems Division,
Hughes Aircraft Company
under

Contract No. NAS8-11421

Space Systems Division
AEROSPACE GROUP
Hughes Aircraft Company · Culver City, California

APPENDIX E

ABSTRACT
FOR
OPTIMAL NONLINEAR FEEDBACK CONTROL DERIVED FROM QUARTIC AND HIGHER-ORDER PERFORMANCE CRITERIA‡

by

R.W. Bass and R.F. Webber

Just as minimization of quadratic performance criteria leads to linear feedback, so it is shown here that minimization of integrals containing quartic or hexadic terms in the state variables leads, respectively, to cubic or quintic feedback. This idea is extended to the minimization of integrals of arbitrarily higher order combinations of the state variables, which is desirable in order to impose inequality constraints upon the state variables. Such laws are shown to be adaptive to actuator saturation (including even bang-bang operation). These results are proved by exhibiting a closed-form solution of the corresponding Hamilton-Jacobi equation, which also provides a globally valid Liapunov function. Prior results of Kalman, Haussler and Rekasius appear as special cases. A new constructive procedure for computing the coefficients of the higher order feedback terms is also presented, together with a numerical application which illustrates remarkable effectiveness in the reduction of overshoots as compared to optimal linear control.

OPTIMAL NONLINEAR FEEDBACK CONTROL DERIVED FROM QUARTIC AND HIGHER-ORDER PERFORMANCE CRITERIA‡

by

R. W. Bass
University of Colorado and Hughes Aircraft Company
Culver City, California, U.S.A.

and

R. F. Webber
Hughes Aircraft Company
Culver City, California, U.S.A.

Introduction

The problem of state-vector feedback control of autonomous, completely controllable linear plants is considered. It seems possible to generalize the following results to multi-channel controllers, but here only a single control variable will be considered. The results derived herein are a natural generalization of results of Kalman[3] for quadratic performance criteria and of results of Haussler[7] and Rekasius[8] for quartic performance criteria.

The present point of view is somewhat different from that of Haussler and Rekasius. Whereas they seek to minimize a quartic criterion subject to a mean-square constraint on the amplitude of the control variable, we impose an additional mean-square-amplitude constraint, namely on what turns out to be the nonlinear feedback part of the control variable. To this problem an exact (not merely "sub-optimal") solution is obtained. By noting that the lower-bound of a nonnegative quantity in the present work is zero, The Haussler-Rekasius quartic upper-bound follows from the present results, while their lower-bound in this context is a consequence of the well-known results on quadratic criteria. Therefore the Haussler-Rekasius results on quartic criteria constitute a genuine corollary of the present completely general results.

The present generalization does not seem to be trivial, however. Firstly, the method[7, 8] of regarding a quartic form of degree n as a sum of $N = (n+2)(n+3)/4!$ squares of quadratic forms seems to us algebraically awkward and more cumbersome to apply numerically than the present technique, as well as unsuited to extension to hexadic and octic forms. Secondly, although the first part of our Theorem 1 can be derived using the Haussler-Rekasius approach[7,8], their measure of "sub-optimality" seems to us unconvincing because for arbitrarily large initial conditions not only the absolute difference $\bar{\Phi} - \underline{\Phi}$ between their upper and lower bounds on the performance criterion Φ (namely $\underline{\Phi} \le \Phi \le \bar{\Phi}$) can be made arbitrarily large, but even the percentage difference $[(\bar{\Phi}/\underline{\Phi}) - 1]$ can be made arbitrarily large; and so the formal reason advanced by them for choosing such a control can be made arbitrarily irrelevant (despite their excellent success in a numerical example[8]). In fact, the striking success of the numerical examples given by Rekasius[8] for n = 2 and by ourselves below for n = 3 seem to be interpretable more conveniently in terms of an optimality attained than a "sub-optimality" which turned out to be better quantitatively than one had any previous, rigorously valid reason to expect.

‡Work performed under NASA Contract NAS 8-11421.

Practical Motivation

This investigation was motivated by a desire to consider the minimax criterion of optimality, namely

$$\min_{\psi} \max_{t} \varphi(x(t)) \qquad (1)$$

where $\varphi(x)$ denotes a positive definite scalar function, x the state vector, t time and ψ the control to be chosen. In practice this criterion may be approximated by the criterion

$$\min_{\psi} \int_0^{+\infty} \left[\varphi\left(x(t)\right)\right]^{\nu} dt \ , \qquad (2)$$

for large integers ν.

Correspondingly, one is led to the general problem of minimizing performance criteria of the type

$$\Phi(x^{\circ}) = \int_0^{+\infty} \Psi\left(x(t)\right) dt \ , \qquad (3)$$

where Ψ is a finite or infinite sum of positive-definite homogeneous multinomial forms of degree 2ν, ($\nu = 1, 2, 3, \cdots$), which constitutes the subject of this paper.

Notation will be established, certain known results reviewed, and certain constructions of multinomial forms defined. Principal results are stated in the form of two theorems, whose proofs are given in Appendix 1. An effective numerical procedure (leading to a computer-oriented system design technique) for finding the required coefficients of higher-order forms is derived in Appendix 2. This design procedure is applied to a third-order numerical example and the results of a computer simulation of the resulting system are presented in Appendix 3. Conclusions follow the main text.

Preliminaries

Vectors are $n \times 1$ columns unless otherwise stated; vector or matrix transposition is denoted by * and scalar product by · ; thus, the Euclidean norm $\|x\|^2 = x^* x = x \cdot x$. Equality by definition is denoted by $\triangleq$.

The performance index to be minimized is of the type $\Phi = \int_0^{+\infty} \Psi\, dt$ defined in (3) above, with

$$\Psi = \sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \Psi_{2\nu} \ , \qquad (4)$$

where

$$\psi_2 \triangleq x \cdot Cx \ , \qquad (C = C^* > 0) \ , \qquad (5)$$

is a given positive-definite homogeneous quadratic form, and where each $\psi_{2\nu} = \psi_{2\nu}(x)$ is a positive semi-definite homogeneous multinomial form of degree 2ν, $(\nu = 2, 3, 4, \cdots)$. (In other words, $\psi_{2\nu}(x) \geq 0$, and $\psi_{2\nu}(\mu x) \equiv \mu^{2\nu}\, \psi_{2\nu}(x)$ for all x and all $\mu \geq 0$.)

The system evolves in time according to

$$\dot{x} = Ax + a\psi \quad , \quad x(0) = x^{\circ} \quad , \quad (^{\bullet} = d/dt) \tag{6}$$

where x is the system state vector, A is the $n \times n$ plant matrix, a is the actuator vector, and ψ is the scalar control law to be chosen in feedback form $\psi = \psi(x)$. It is assumed throughout that (6) is controllable,[3] in the sense that the vectors $A^{i-1}a$, $(i = 1, 2, \cdots, n)$, are linearly independent. Control laws are admissable only if they produce asymptotic stability of the equilibrium state $x = 0$; in particular, it is required that

$$x(t) \to 0 \quad \text{as} \quad t \to +\infty \tag{7}$$

This stability will be established by explicit construction of a Liapunov function $V = V(x)$, of the form

$$V \triangleq \sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) \varphi_{2\nu} \quad , \tag{8}$$

where

$$\varphi_2 \triangleq x \cdot Bx \quad , \quad (B = B^* > 0) \quad , \tag{9}$$

is a positive-definite quadratic form, where each $\varphi_{2\nu} = \varphi_{2\nu}(x)$ is a positive semi-definite homogeneous multinomial form of degree 2ν, $(\nu = 2, 3, 4, \cdots)$, and where Liapunov's stability theory[1, 2] is applicable by virtue of the fact that the Lie derivative of $V(x)$ along the vector field (6) is a negative-definite function, namely $-\Psi(x) - (Ba \cdot x)^2$. In other words, $\psi = \psi(x)$ will be so chosen that whenever $x = x(t)$ satisfies (6),

$$\dot{V}\left(x(t)\right) \equiv -\Psi\left(x(t)\right) - [g \cdot x(t)]^2 \quad , \tag{10}$$

where

$$g \triangleq -Ba \quad . \tag{11}$$

Note that

$$\text{grad}_{(x)} V(x) = Bx + \sum_{\nu=2}^{\infty} \left(\frac{1}{2\nu}\right) \text{grad}\, \varphi_{2\nu}(x) \quad , \tag{12}$$

whence, using the definitions (11) and

$$\psi_{n\ell} \triangleq - \sum_{\nu=2}^{\infty} \left(\frac{1}{2\nu}\right) a \cdot \text{grad}\, \varphi_{2\nu} \quad , \tag{13}$$

it is clear that the scalar quantity

$$\sigma \triangleq -a \cdot \text{grad}\, V \equiv - \sum_{\nu=1}^{\infty} \left(\frac{1}{2\nu}\right) a \cdot \text{grad}\, \varphi_{2\nu} \tag{14}$$

can be expressed as

$$\sigma = g \cdot x + \psi_{n\ell}(x) \quad . \tag{15}$$

The quantities σ and $\psi_{n\ell}$ are important in forming the optimal control law ψ, and the definitions (13)-(14) and identity (15) will be assumed henceforth and used repeatedly in the sequel without further reference.

To recapitulate, the matrix-vector pair (A, a) and the functions $\{\psi_{2\nu}\}$ are given, while the functions $\{\varphi_{2\nu}\}$ are to be constructed by algebraic operations upon (A, a) and the coefficients of $\{\psi_{2\nu}\}$; then $V(x) = \sum_{\nu=1}^{\infty} (1/2\nu)\, \varphi_{2\nu}(x)$ is defined by (8), and the functions $\sigma(x)$ and $\psi_{n\ell}(x)$ are to be found from the definitions (9), (11)-(15).

The method of computing the $\{\varphi_{2\nu}\}$ from the $\{\psi_{2\nu}\}$ will be prefaced by a special case, namely computation of $\varphi_2 = x \cdot Bx$ from $\psi_2 = x \cdot Cx$. This in turn will be motivated by reviewing the well-known results of Kalman[3] regarding linear regulators.

Quadratic Performance Criteria

Consider now the case where

$$\Psi = \frac{1}{2}\psi_2 = \frac{1}{2} x \cdot Cx \quad , \tag{16}$$

and choose ψ so as to minimize (3) subject to a "mean-square amplitude" constraint on ψ of the type

$$\frac{1}{2}\int_0^{+\infty} \psi^2\, dt \leq \rho_1 = \text{const.} \tag{17}$$

By the Lagrange multiplier technique [absorb the multiplier into C, by allowing C to be multiplied by any positive scalar without loss of generality] the minimization of (3) and (16), subject to the constraint (17), can be replaced by the unconstrained minimization over ψ of

$$\Phi = \Phi(x^{\circ}) = \frac{1}{2}\int_0^{+\infty} [x \cdot Cx + \psi^2]\, dt \quad . \tag{18}$$

It is well known[3] that the solution to the problem of choosing ψ in (6) to minimize the Φ of (18) is given by

$$\psi = g \cdot x \quad , \quad (g = -Ba) \quad , \tag{19a}$$

$$\min \Phi = \Phi(x) = \frac{1}{2} x \cdot Bx \quad , \tag{19b}$$

where $B = B^* > 0$ is the unique positive definite solution of the (equilibrium) matrix Riccati equation

$$BA + A^*B - Baa^*B = -C \quad . \tag{20}$$

Inserting (19) into (6) displays the controlled system in its "closed-loop form" as a linear system

$$\dot{x} = Ax + a\psi = \tilde{A}x \quad , \tag{21}$$

where by definition

$$\tilde{A} \triangleq A + ag^* \equiv A - aa^*B \quad , \tag{22}$$

and where $\tilde{A}$ is known to be a stability (Hurwitz) matrix. On subtracting $-Baa^*B \equiv -gg^*$ from both sides of (20), and defining the positive-definite matrix $\tilde{C}$ by

$$\tilde{C} \triangleq C + gg^* \quad , \tag{23}$$

equation (20) becomes

$$B\tilde{A} + \tilde{A}^*B = -\tilde{C} \quad , \quad (\tilde{C} = \tilde{C}^* > 0) \quad , \tag{24}$$

which has the well-known solution[5]

$$B = B^* = \int_0^{+\infty} \exp(\tilde{A}^*t)\, \tilde{C}\, \exp(\tilde{A}t)\, dt > 0 \quad . \tag{25}$$

A highly practical, purely algebraic algorithm for computing g (without first finding B) is given by Bass and Gura[4]; from g, both $\tilde{A}$ and $\tilde{C}$ can be found readily, and then B can be computed from

(24) either by the inversion of a matrix of order $1/2\ n(n+1)$ as in Bellman[5] (p. 231), or by the technique of Appendix 2 below. Alternatively, a purely algebraic algorithm for computing B directly from (a, A, C) which involves multiplying a $2n \times 2n$ matrix by itself (n-1) times and then inverting a matrix of order n is given by Bass[9].

Note that, on recalling $\varphi_2 = x \cdot Bx$ and defining

$$\tilde{\psi}_2 \triangleq \psi_2(x) + (g \cdot x)^2 \quad , \quad g = -Ba \quad , \tag{26}$$

the algebraic equation (24) takes the form [just premultiply by x^* and postmultiply by x] of the partial differential equation

$$\tilde{A}x \cdot \text{grad}\ \varphi_2(x) = -\tilde{\psi}_2(x) \quad . \tag{27}$$

A Theorem of Liapunov

Equation (27) illustrates a classic theorem of Liapunov[1, 2], which shows that if $\tilde{A}$ is an arbitrary stability matrix, and if $\psi_{2\nu}(x)$ is any positive semi-definite homogeneous multinomial form of degree 2ν, the partial differential equation

$$\tilde{A}x \cdot \text{grad}\ \varphi_{2\nu}(x) = -\psi_{2\nu}(x) \quad , \quad (\nu = 1, 2, 3, \cdots) \quad , \tag{28}$$

has a unique solution $\varphi_{2\nu}(x)$ which is also a positive semi-definite homogeneous multinomial form of degree 2ν. A new practical algorithm for solving (28) is given in Appendix 2. Henceforth it will be assumed that the $\{\varphi_{2\nu}\}$ are constructed from the $\{\psi_{2\nu}\}$ in accordance with (28), for $\nu = 2, 3, 4 \cdots$.

Summary of Algebraic Constructions

To recapitulate, the pair (A, a) is given together with the forms $\{\psi_{2\nu}\}$. First $B=B(A, a, C)$ is constructed so that $\tilde{A} = A - aa^*B$ is a stability matrix satisfying (27) with

$$\varphi_2 = x \cdot Bx \quad , \quad \tilde{\psi}_2 = x \cdot Cx + (g \cdot x)^2 \quad . \tag{29}$$

Then this $\tilde{A}$ and the $\{\psi_{2\nu};\ \nu = 2, 3, 4, \cdots\}$ are used to construct the remaining $\{\varphi_{2\nu};\ \nu=2, 3, 4, \cdots\}$ so that (28) holds. Now $V(x)$, $\sigma(x)$, and $\psi_{n\ell}(x)$ can be constructed as in (8), (13), (14), and will henceforth be regarded as known quantities.

Principal Results

Consider the choice of ψ in (6) to effect minimization of the general Φ of (3), subject to the constraint (17) and an additional constraint of the type

$$\frac{1}{2}\int_0^{+\infty} [\psi_{n\ell}(x(t))]^2\, dt \le \rho_2 = \text{const.} \tag{30}$$

The constraint (30) is at this stage admittedly a somewhat artificial condition, justified only because it permits an explicit, closed-form solution of the problem at hand. However, it will turn out a posteriori that $\psi_{n\ell}(x)$ happens to agree with the nonlinear terms in the optimal feedback control law $\psi(x)$; hence the physical meaning of the two independent constraints (17) and (13) is that the "mean-square-amplitudes" of both the linear and the nonlinear terms in the optimal control law must be a priori bounded separately.

Once again, the Lagrange multiplier technique may be used to formulate an equivalent unconstrained problem, namely, that of choosing the control law ψ in (6) so as to minimize the unconstrained performance criterion

$$\Phi = \Phi(x^\circ) \triangleq \int_0^{+\infty} \left\{ \chi(x) + \frac{1}{2}\psi^2 + \frac{1}{2}[\psi_{n\ell}(x)]^2 \right\} dt \, . \tag{31}$$

It is important to note that ρ_1 in (17) and ρ_2 in (30) can be chosen independently and arbitrarily. At first glance this seems to require an integrand in (31) of the form $\chi + (1/2)\lambda_1\psi^2 + (1/2)\lambda_2\psi_{n\ell}^2$. However, on replacing ψ_2 by $\lambda_1\psi_2$, and $\psi_{2\nu}$ by $\sqrt{\lambda_1/\lambda_2}\,\psi_{2\nu}$ for $\nu \ge 2$, the quantity $\lambda_2\psi_{n\ell}^2$ is replaced by $\lambda_1\psi_{n\ell}^2$. Hence division by λ_1 yields an integrand of the form $\lambda_1 = \lambda_2 = 1$, in which now each $\psi_{2\nu}$, $\nu \ge 2$, has been replaced by $(1/\sqrt{\lambda_1\lambda_2})\,\psi_{2\nu}$. Thus by letting scalar factors multiplying ψ_2 and $\psi_{2\nu}$, $\nu \ge 2$, run independently over all positive values, all constraints ρ_1 and ρ_2 will be attained. (In numerical applications of (18) it is well known that a factor multiplying C must be allowed to vary similarly in order to insure attainment of (17).)

Theorem 1

The optimal control law for (6) relative to (31) is given by

$$\psi = \sigma(x) \equiv g \cdot x + \psi_{n\ell}(x) \quad , \tag{32}$$

and, furthermore,

$$V(x^\circ) = \min_{\psi} \Phi(x^\circ) \quad . \tag{33}$$

Moreover, the related control law

$$\psi = \mu\sigma \tag{34}$$

yields global asymptotic stability of $x = 0$ for all μ such that

$$\mu > \frac{1}{2} \quad . \tag{35}$$

Theorem 2

Let

$$\mu > \frac{1}{2} \, , \quad \varkappa > 0 \quad , \tag{36}$$

be arbitrary numbers. Choose $\varepsilon = \varepsilon(\varkappa) > 0$ so small that on the neighborhood of $x = 0$ defined by

$$V(x) < \varepsilon \tag{37}$$

the inequality

$$|\sigma(x)| < 2\varkappa \tag{38}$$

holds everywhere. Then the control law

$$\psi = \varkappa\ \text{sat}\,[\mu\sigma/\varkappa]$$

yields asymptotic stability of $x = 0$ on the region (37).

The practical utility of the preceding results may be inferred from the application summarized in Appendix 3.

Conclusions

A completely general algorithm has been presented whereby nonlinear feedback laws can be computed which minimize integral performance criteria defined by multinomial forms of higher than quadratic order. A criterion of order 2ν yields a feedback control law of order $2\nu - 1$, ($\nu = 1, 2, 3, \cdots$). These results represent a generalization of the results of Kalman[3], Haussler[7], and Rekasius[8].

Minimax criteria can be approximated more and more closely by increasing ν; however, it does not seem practical to take ν very large, because there will in general be $N = n(n+1) \cdots (n + 2\nu - 1)/(2\nu)!$ distinct nonlinear feedback terms which must be mechanized.

Practical experience to date indicates very satisfactory results with $\nu = 2$. That is, quartic criteria will keep the state variables (or linear combinations thereof) very nearly within pre-specified allowable bounds, while the required cubic feedback control law is feasible to mechanize.

Appendix 1

Proof of Theorem 1

The law $\psi = \sigma$ provides a unique solution to the Hamilton-Jacobi equation

$$\mathcal{K}(x, y, \psi) = \overline{\mathcal{K}}(x, y) = 0 \quad , \tag{39}$$

$$\overline{\mathcal{K}}(x, y) \stackrel{\Delta}{=} \max_{\psi} \mathcal{K}(x, y, \psi) \ , \quad y \stackrel{\Delta}{=} -\text{grad}\ \Phi \tag{40}$$

In fact, taking

$$\mathcal{K} \stackrel{\Delta}{=} y \cdot Ax + (a \cdot y)\ \psi - \Psi - \frac{1}{2}\psi^2 - \frac{1}{2}\psi_{n\ell}^{\ 2} \quad , \tag{41}$$

and noting that $\partial\mathcal{K}/\partial\psi \equiv a \cdot y - \psi = 0$ if and only if

$$\psi = a \cdot y = -\, a \cdot \text{grad}\ \Phi \quad , \tag{42}$$

while $\partial^2\mathcal{K}/\partial\psi^2 = -1 < 0$ at $\psi = a \cdot y$, one obtains

$$\overline{\mathcal{K}}(x, y) = \mathcal{K}(x, y, a \cdot y) \quad . \tag{43}$$

Hence (39) becomes, by (42),

$$\begin{aligned}
\mathcal{K} &= y \cdot Ax + \frac{1}{2}\psi^2 - \frac{1}{2}[\psi - g \cdot x]^2 - \Psi = \\
&= y \cdot \widetilde{A}x - \psi(g \cdot x) + \frac{1}{2}\psi^2 - \frac{1}{2}\psi^2 + \psi(g \cdot x) + \\
&\quad - \frac{1}{2}(g \cdot x)^2 - \Psi = \\
&= y \cdot \widetilde{A}x - \Psi - \frac{1}{2}(g \cdot x)^2 = \\
&= -\left\{\frac{1}{2}[\widetilde{A}x \cdot \text{grad}\ \varphi_2 + \widetilde{\psi}_2] + \sum_{\nu=2}^{\infty}\left(\frac{1}{2\nu}\right)[\widetilde{A}x \cdot \text{grad}\varphi_{2\nu} + \psi_{2\nu}]\right\} \\
&\equiv 0
\end{aligned} \tag{44}$$

by (27)-(28), provided that it is possible to identify V and Φ and so use

$$y = -\text{grad}\ V = -\sum_{\nu=1}^{\infty}\left(\frac{1}{2\nu}\right) \text{grad}\ \varphi_{2\nu} \quad , \tag{45a}$$

$$\psi = a \cdot y \equiv \sigma \equiv g \cdot x + \psi_{n\ell} \quad . \tag{45b}$$

However, comparing (41) and (44), and using (45), (39) may be expressed as

$$\dot{V} \equiv \dot{x} \cdot \text{grad}\ V = -\left[\Psi + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^{\ 2}\right] \quad . \tag{46}$$

Thus, by Liapunov's direct method, $x(t) \to 0$ as $t \to +\infty$, and, integrating (46), one obtains the result that $\psi = \sigma$ implies $V(x^\circ) = \Phi(x^\circ)$.

Similarly, upon choosing $\psi = \mu\sigma$, it can be shown that

$$\dot{V} = -\left[\Psi + \frac{1}{2}\psi_{n\ell}^{\ 2}\right] - \left(\mu - \frac{1}{2}\right)\sigma^2 \quad . \tag{47}$$

This completes the proof of Theorem 1.

Proof of Theorem 2

Assume $V(x) < \epsilon = \epsilon(\varkappa)$ and $|\sigma(x)| < 2\varkappa$. Then it can be shown easily, using V(x) as defined by (8) and $\psi = \varkappa\ \text{sat}\ (\mu\,\sigma(x)/\varkappa)$, that

$$\dot{V}(x) = -\Psi(x) + \frac{1}{2}\sigma^2(x) - \varkappa\sigma(x)\text{sat}\left[\frac{\mu}{\varkappa}\sigma(x)\right] - \frac{1}{2}\psi_{n\ell}^{\ 2}(x) \quad . \tag{48}$$

Now consider two cases, (i) and (ii). For (i) let $|(\mu/\varkappa)\,\sigma(x)| \le 1$ and for (ii) let $|(\mu/\varkappa)\,\sigma(x)| \ge 1$. Then for (i),

$$\dot{V} = -\Psi - \frac{1}{2}\psi_{n\ell}^{\ 2} - \left(\mu - \frac{1}{2}\right)\sigma^2 \quad , \tag{49}$$

and $\dot{V} < 0$ when $\mu > \frac{1}{2}$.

For (ii),

$$\dot{V} = -\Psi - \frac{1}{2}\psi_{n\ell}^{\ 2} + \frac{1}{2}\sigma^2 - \varkappa|\sigma| \quad , \tag{50}$$

and $\dot{V} < 0$ for $|\sigma| < 2\varkappa$. This completes the proof of Theorem 2.

Appendix 2

A Theory of Higher Order Forms

Presented here are techniques for effective use of $2\nu^{\text{th}}$ order forms. As explained following (28), construction of an optimal control depends on solving the equation

$$\widetilde{A}x \cdot \text{grad}\ \varphi_{2\nu}(x) = -\psi_{2\nu}(x) \quad , \tag{51}$$

for $\varphi_{2\nu}(x)$. This relation actually represents N linear equations in N unknowns. It will be seen in the sequel to (58) below that the dimension N is

$$N = \frac{n(n+1)\cdots(n+2\nu-1)}{(2\nu)!} \tag{52}$$

The unknowns are the coefficients of the different terms in the $2\nu^{\text{th}}$ order form $\varphi_{2\nu}(x)$ and the knowns are the corresponding coefficients in $\psi_{2\nu}(x)$. Thus (28) may be represented by

$$\mathcal{A}b = c \tag{53a}$$

where

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_N \end{pmatrix} , \quad c = \begin{pmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ c_N \end{pmatrix} , \tag{53b}$$

and $\mathcal{A}$ is an $N \times N$ matrix. The b_i's represent coefficients in the unknown $\varphi_{2\nu}(x)$ and the c_i's represent the corresponding coefficients in the known $\psi_{2\nu}(x)$. In order to solve for b it is necessary, in effect, to invert $\mathcal{A}$. This could be accomplished by standard techniques. That is, just write out the relationships involved and solve for the b_i's. This however would require a considerable amount of algebra even for simple problems. For example, if $n = 5$ and $2\nu = 4$, then $N = 70$.

As an alternative to solving (53a) in this manner, one might use the spectral resolution of $\mathcal{A}$. Thus, seek the eigenfunctions of the operator.

$$\widetilde{A}x \cdot \text{grad}_{(x)}(\cdot) \quad .$$

In other words, seek homogeneous multinomial forms $\zeta_k(x)$ of degree 2ν such that, for (complex) eigenvalues μ_k,

$$\widetilde{A}x \cdot \text{grad}\ \zeta_k(x) = \mu_k\ \zeta_k(x) \quad . \tag{54}$$

It will be shown that the ζ_k's can be formed from products of linear forms raised to various integer powers. This idea will now be presented in detail.

Let u^k be right eigenvectors of $\widetilde{A}$, let v^k be left eigenvectors of $\widetilde{A}$, and let the corresponding eigenvalues be λ_k. Then

$$\widetilde{A}u^k = \lambda_k u^k \quad , \tag{55a}$$

$$\widetilde{A}^* v^k = \lambda_k v^k \quad . \tag{55b}$$

It will be assumed that the λ_k are distinct $(k=1, 2, \cdots, n)$; then the $\{u^k\}$, $\{v^k\}$ are linearly independent and can be so normalized that

$$u^k \cdot v^j = \delta_{kj} \quad , \quad (k, j = 1, \cdots, N) \quad . \tag{56}$$

Once the λ_k's are known, the calculation of the normalized u^k's and v^k's may be easily accomplished, e.g., using the closed form expressions presented by Bass and Gura[6].

Define $\zeta_\ell(x)$ by

$$\zeta_\ell(x) \triangleq (v^1 \cdot x)^{m_{1\ell}} (v^2 \cdot x)^{m_{2\ell}} \cdots (v^n \cdot x)^{m_{n\ell}} \quad , \tag{57}$$

where the $m_{i\ell}$'s are integers determined by

$$\sum_{i=1}^{n} m_{i\ell} = 2\nu \quad , \quad m_{i\ell} \geq 0 \quad . \tag{58}$$

The expression (58) does not uniquely determine the $m_{i\ell}$'s. Therefore, let ℓ be an index corresponding to each permissible set $\{m_{i\ell}\}$. It is shown in Malkin[2] that there are N such sets, where N is given by (52).

Using (57) straightforward manipulations yield

$$\widetilde{A}x \cdot \mathrm{grad}\, \zeta_\ell(x) \equiv \widetilde{A}x \cdot \mathrm{grad}\left[(v^1 \cdot x)^{m_{1\ell}} \cdots (v^n \cdot x)^{m_{n\ell}}\right] =$$

$$= (\widetilde{A}x \cdot v^1) m_{1\ell} (v^1 \cdot x)^{m_{1\ell}-1} \cdots (v^n \cdot x)^{m_{n\ell}} +$$

$$+ (\widetilde{A}x \cdot v^2) m_{2\ell} (v^1 \cdot x)^{m_{1\ell}} (v^2 \cdot x)^{m_{2\ell}-1} \cdots (v^n \cdot x)^{m_{n\ell}} + \cdots +$$

$$+ (\widetilde{A}x \cdot v^n) m_{n\ell} (v^1 \cdot x)^{m_{1\ell}} \cdots (v^n \cdot x)^{m_{n\ell}-1} =$$

$$= (m_{1\ell}\lambda_1 + m_{2\ell}\lambda_2 + \cdots + m_{n\ell}\lambda_n)(v^1 \cdot x)^{m_{1\ell}} \cdots (v^n \cdot x)^{m_{n\ell}} =$$

$$= \mu_\ell \zeta_\ell(x) \quad . \tag{59a}$$

Malkin[2] has shown that by letting the $m_{i\ell}$'s range over all permissible values, as given by (58), one does in fact, exhaust all the eigenvalues $\{\mu_\ell\}$ of the operator $\widetilde{A}x \cdot \mathrm{grad}_{(x)}(\cdot)$. If it is effectively possible to expand $-\psi_{2\nu}(x)$ in the eigenfunctions $\{\zeta_\ell\}$ then the equation (51) can be solved for $\varphi_{2\nu}(x)$ by identifying coefficients in eigen-expansions. Specifically, if

$$-\psi_{2\nu}(x) = \sum_{\ell=1}^{N} \gamma_\ell \zeta_\ell(x) \quad , \tag{59b}$$

then

$$\varphi_{2\nu}(x) = \sum_{\ell=1}^{N} (\gamma_\ell/\mu_\ell)\, \zeta_\ell(x) \quad . \tag{59c}$$

Begin by assuming, for a typical term of $-\psi_{2\nu}(x)$,

$$\lambda x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n} \quad , \quad \lambda = \text{constant}. \tag{60}$$

Expand each term (60) in eigenfunctions as follows. Write for $x_k^{m_k}$

$$x_k^{m_k} \triangleq (e^k \cdot x)^{m_k} \tag{61}$$

and then expand e^k in terms of the left eigenvectors of $\widetilde{A}$. That is, expand e^k as

$$e^k = \sum_{i=1}^{n} \alpha_{ik} v^i \quad , \quad (k = 1, \cdots, n) \quad . \tag{62}$$

From the theory of matrices one has

$$I_n = \sum_{i=1}^{n} v^i (u^i)^* \quad ,$$

whence

$$\alpha_{ik} = (u^i \cdot e^k) \quad , \quad (i, k = 1, \cdots, n) \quad . \tag{63}$$

Thus, by using (61), (62), and (63), $x_k^{m_k}$ may be expressed as

$$x_k^{m_k} = (e^k \cdot x)^{m_k} = \left\{\sum_{i=1}^{n} (u^i \cdot e^k)(v^i \cdot x)\right\}^{m_k} \quad . \tag{64}$$

Recall now that, by (57), the $(v^i \cdot x)$ are the linear forms used to obtain $\zeta_\ell(x)$. Thus, when (64) is put into (60) and multiplied out, there results the desired expansion

$$-\psi_{2\nu}(x) = \sum_{\ell=1}^{N} \beta_\ell \zeta_\ell(x) \quad . \tag{65}$$

In practice, expansion of (60), though straightforward in nature, requires considerable symbolic multiplication of multinomials.

A Numerical Example

As an example of the foregoing procedure, consider the following case wherein $n=2$. Let $\widetilde{A} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$; then the eigenvalues of $\widetilde{A}^*$ are $-2, -1$; and the corresponding eigenvectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Now let it be desired to solve the equation $Ax \cdot \mathrm{grad}\, \varphi_4(x) = -\psi_4(x)$ and let $\psi_4(x) = (x_1 + x_2)^4$. Then the ζ_k's and μ_k's are given by

$$\zeta_1 = (x_1 + x_2)^4 \quad , \quad \mu_1 = -8 = (-2)4 + (-1)(0) \quad ,$$

$$\zeta_2 = (x_1 + x_2)^3 (2x_1 + x_2) \quad , \quad \mu_2 = -7 \quad ,$$

$$\zeta_3 = (x_1 + x_2)^2 (2x_1 + x_2)^2 \quad , \quad \mu_3 = -6 \quad ,$$

$$\zeta_4 = (x_1 + x_2)(2x_1 + x_2)^3 \quad , \quad \mu_4 = -5 \quad ,$$

$$\zeta_5 = (2x_1 + x_2)^4 \quad , \quad \mu_5 = -4 \quad .$$

Solving for the coefficients in the eigen-expansion of $\varphi_4(x)$ yields

$$\varphi_4(x) = \frac{1}{8}(x_1 + x_2)^4 \quad .$$

An Alternative Procedure

In equations (57), (60)-(64), and (65) a general method was presented for expanding

$-\psi_2(x)$ in eigenfunctions. Sometimes it is easier to accomplish this expansion directly without recourse to the right eigenvectors $\{u^k\}$ of $\tilde{A}$, which are needed if (63) is used. Again an example provides a convenient means of illustrating this. Let

$$\psi_4(x) = x_1^2 x_2^2 + x_2^4 ,$$

and let the eigenfunctions be the same as in the previous example. Then it is only necessary to write x_1 and x_2 in terms of the linear forms $(x_1 + x_2)$ and $(2x_1 + x_2)$. The proper expansions are

$$x_1 = (2x_1 + x_2) - (x_1 + x_2) , \quad x_2 = 2(x_1 + x_2) - (2x_1 + x_2) .$$

Define

$$x_1 + x_2 \triangleq \alpha , \quad 2x_1 + x_2 \triangleq \beta .$$

Then

$$\psi_4(x) = (\beta - \alpha)^2(2\alpha - \beta)^2 + (2\alpha - \beta)^4 = -\sum_{\ell=1}^{5} \gamma_\ell \zeta_\ell .$$

Expanding this expression in α, β and noting that

$$\zeta_1 = \alpha^4 , \quad \zeta_4 = \alpha^3 \beta$$

$$\zeta_2 = \alpha^2 \beta^2 , \quad \zeta_5 = \alpha\beta^3$$

$$\zeta_3 = \beta^4$$

yields the desired coefficients $\gamma_1, \gamma_2, \cdots, \gamma_5$.

Appendix 3

Simulated Example of Stability Augmentation by Cubic Feedback

Consider the system

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = x_3$$

$$\dot{x}_3 = -6x_1 - 11x_2 - 6x_3 + \psi_{n\ell}(x) .$$

If we let x_1 represent position, then x_2 and x_3 represent velocity and acceleration, respectively. The object is to choose the feedback control $\psi_{n\ell}(x)$ so that large overshoots in velocity or acceleration are avoided when the initial displacement is $x_1(0) = x_0$, $x_2(0) = 0$ and $x_3(0) = 0$.

As the system returns to the origin x_2 (velocity) or x_3 (acceleration) may be prohibitively large. It is necessary to apply nonlinear feedback in an appropriate manner to reduce the offending state.

To accomplish this, we consider the performance indexes Φ_1 and Φ_2 where

$$\Phi_1 = \int \left[x_2^4 + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^2 \right] dt$$

and

$$\Phi_2 = \int \left[x_3^4 + \frac{1}{2}\psi^2 + \frac{1}{2}\psi_{n\ell}^2 \right] dt .$$

Minimization of Φ_1 or Φ_2 can be effected by cubic feedback $\psi_{n\ell}$, where $\psi_{n\ell}(\mu x) \equiv \mu^3 \psi_{n\ell}(x)$, and where $\psi_{n\ell}$ is defined by Theorem 1 and is computable as in Appendix 2.

The feedback control $\psi_{n\ell}$ derived from Φ_1 will keep x_2 small, and the control $\psi_{n\ell}$ derived from Φ_2 will keep x_3 small. In Figure 1 the phase-plane plot of x_2 versus x_3 is shown. Included in this figure is the response of the stable linear system. The initial conditions for the responses are $x(0) = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix}$ and $x(0) = \begin{pmatrix} 20 \\ 0 \\ 0 \end{pmatrix}$.

By examining this plot the reduction in overshoot of x_2 and x_3 becomes apparent.

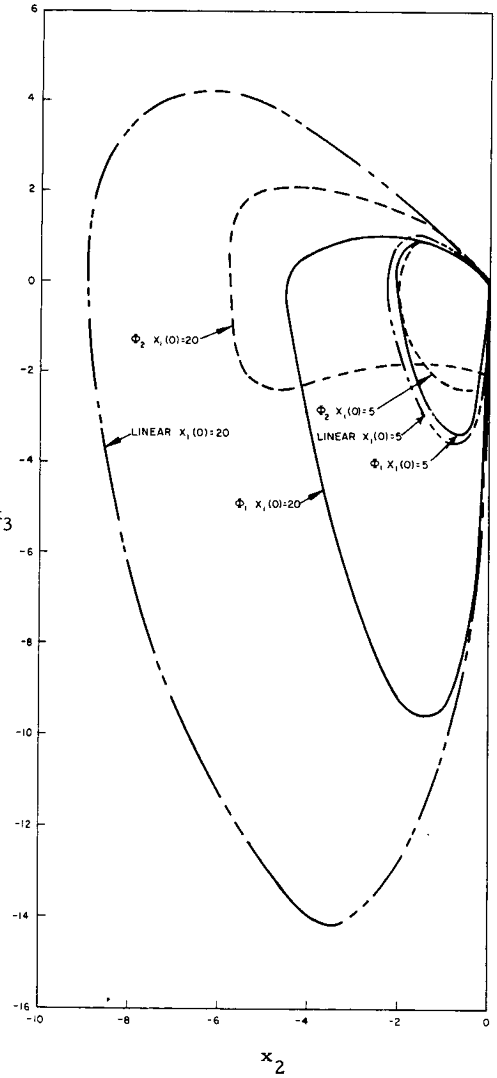


Figure 1. Linear and Non-linear Feedback Comparison

References

1. A. Liapunov, Problème Generale de la Stabilité du Mouvement, Princeton University Press, 1949, pp. 276-277.

2. I. Malkin, Theory of Stability of Motion, Moscow, 19 (English translation available from Office of Technical Services, U.S. Atomic Energy Commission).

3. R.E. Kalman, "When is a Linear Control Optimal?," Journal of Basic Engineering, March 1964, pp. 51-60.

4. R.W. Bass and I. Gura, "High Order System Design via State Space Considerations," 1965 JACC Proceedings, pp. 311-318.

5. R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.

6. R.W. Bass and I. Gura, "Canonical Forms for Controllable Systems with Applications to Optimal Nonlinear Feedback," (Submitted to 1966 IFAC).

7. R.L. Haussler and Z.V. Rekasius, "Über die sub-optimale regelung von nichtlinearen systemen," Regelungstechnik, July 1964.

8. Z. Rekasius, "Sub-Optimal Design of Intentionally Non-Linear Systems," 1964 JACC Proceedings, pp. 436-440, also IEEE Transactions on Automatic Control, October 1964, Vol. AC-9, No. 4, pp. 380-385.

9. R.W. Bass, "Solution of the Matrix Riccati Equation by Spectral Factorization," (to appear).

APPENDIX F

CONTROLLABILITY

by

I. Gura
Space Systems Division,
Hughes Aircraft Company
under

Contract No. NAS8-11421

Space Systems Division
Guidance and Controls Division
AEROSPACE GROUP
Hughes Aircraft Company · Culver City, California

APPENDIX F

CONTROLLABILITY

DEFINITION

The system

$$\dot{x} = Ax + a\psi \ , \qquad x(0) = x^o \tag{F-1}$$

is said to be controllable if for every initial state x^o, there is a control law $\psi = \psi(t)$ such that the solution $x(t)$ of (F-1) satisfies $x(T) = x^1$ where x^1 is arbitrary and $T > 0$.

<u>Theorem</u>. A necessary and sufficient condition for the system (F-1) to be controllable is that

$$\det(a, Aa, \cdots, A^{n-1}a) = \det(D) \neq 0 \tag{F-2}$$

<u>Proof</u>

<u>Part I — Necessity</u>. (If $\det(D) = 0$ there is an x^1 such that no control law $\psi(t)$ can transfer the system (F-1) from some x^o to x^1.)

In general, the solution of (F-1) is

$$x(t) = \exp\,(At)x^o + \int_o^t \exp\left[A(t - \lambda)\right] a\psi(\lambda)d\lambda \tag{F-3}$$

At $x(T) = x^1$

$$x^1 - \exp\,(AT)x^o = \int_o^T \exp\left[A(T - \lambda)\right] a\psi(\lambda)d\lambda \tag{F-4}$$

Now since

$$\exp(A\lambda) = \sum_{k=0}^{\infty} \frac{A^k \lambda^k}{k!} \tag{F-5}$$

(F-4) becomes

$$\exp(-AT)x^1 - x^o - \left\{ \sum_{k=0}^{\infty} A^k a \int_o^T \frac{\lambda^k}{k!} \psi(\lambda) d\lambda \right\} \tag{F-6}$$

Now if

$$\det(a, Aa, \cdots, A^{n-1}a) = 0 \quad , \tag{F-7}$$

the system of equations

$$(a, Aa, \cdots, A^{n-1}a)^* q = 0 \tag{F-8}$$

has a solution $q \neq 0$. Thus there is a vector q such that

$$q \cdot A^{i-1} a = 0 \qquad (i = 1, 2, \cdots) \tag{F-9}$$

Note that (F-9) is valid for all $i > n$ since by the Cayley-Haminton Theorem $A^j (j \geq n)$ can always be expressed as a polynomial in powers of A less than $n - 1$.

By (F-6)

$$q \cdot \left(\exp(-AT)x' - x^o \right) = 0 \tag{F-10}$$

But this is incompatible with the choice

$$x^1 = \exp(AT)x^o + q \tag{F-11}$$

since (F-10) would then imply

$$q \cdot q = 0 \qquad q \neq 0 \quad . \tag{F-12}$$

Thus the system (F-1) <u>cannot</u> be controllable if $\det(D) = 0$.

<u>Part II — Sufficiency</u>. (If $\det(D) \neq 0$ there is always a way of picking a control to transfer the system (F-1) from x^o to x^1).

Choose

$$\psi(t) = a \cdot y(t) \tag{F-13}$$

where $y(t)$ is defined by the solution of

$$y = -A^* y \ , \qquad y(0) = y^o \tag{F-14}$$

Clearly

$$y = \exp(-A^* t) y^o \tag{F-15}$$

and so

$$\psi(t) = a \cdot y(t) = \left[\exp(-AT)a\right] \cdot y^o \tag{F-16}$$

Thus (F-4) becomes

$$\exp(-AT)x^1 - x^o = \left\{\int_o^T \exp(-A\lambda)aa^* \exp(-A^*\lambda)d\lambda\right\} y^o \tag{F-17}$$

Now define a matrix P by

$$P = \int_o^T \exp(-A\lambda)aa^* \exp(-A^*\lambda)d\lambda \tag{F-18}$$

Then (F-17) becomes

$$Py^o = \exp(-AT)x^1 - x^o \tag{F-19}$$

If $\det(P) \neq 0$, the desired control law described in (F-13)-(F-14) will be completely determined since then

$$y^o = P^{-1}\left|\exp(-AT)x^1 - x^o\right| \tag{F-20}$$

To consider this possibility, note from (F-18) that

$$z \cdot Pz = \int_o^T \left(a \cdot \exp(-A^*\lambda)z^{\,2}\right)d\lambda \tag{F-21}$$

Obviously, if

$$a \cdot \exp(-A^*z) \not\equiv 0 \tag{F-22}$$

then $z \cdot Pz > 0$ and P must be invertible. (The determinant of a matrix is equal to the product of its eigenvalues and since P must have positive eigenvalues $\det(P) \neq 0$.)

Assume that

$$a \cdot \exp(-A^*\lambda)z \equiv 0 \ , \qquad z \neq 0 \tag{F-23}$$

Then by repeated differentiations with respect to λ

$$(A^{j-1}a) \cdot \exp(-A^*\lambda)z \equiv 0 \ , \qquad (j = 1, 2, \cdots, n) \tag{F-24}$$

Now since $\left|\exp(-A^*\lambda)\right|^{-1} = \exp(A^*\lambda)$ always exists, $\exp(-A^*\lambda)z = 0$ can only be valid of $z = 0$. Since this is ruled out by hypothesis, $\exp(-A^*\lambda)z \neq 0$ and (F-24) can hold only if

$$\det(a, Aa, \cdots, A^{n-1}a) = 0 \ . \tag{F-25}$$

Thus if (F-25) is ruled out, $P > 0$ and (F-20) is valid, thus proving the theorem.

OBSERVABILITY

Theorem. If

$$\dot{x} = Ax \; , \qquad x(0) = x^{o} \neq 0 \tag{F-26}$$

$$\operatorname{rank}\left[H^{*},\ A^{*}H^{*},\ (A^{*})^{2} H^{*},\ \ldots,\ (A^{*})^{n-1} H^{*}\right] = n \tag{F-27}$$

Then

$$\|Hx\|^{2} \not\equiv 0 \tag{F-28}$$

Proof. Assume

$$\|Hx\| = 0 \; , \qquad x^{o} \neq 0 \tag{F-29}$$

This implies that

$$\begin{aligned} Hx &\equiv 0 \\ H \frac{dx}{dt} &\equiv 0 \\ H \frac{dx^{2}}{dt^{2}} &\equiv 0 \\ \cdots&\cdots \\ H \frac{d^{n-1}x}{dt^{n-1}} &\equiv 0 \end{aligned} \tag{F-30}$$

However, since the solution of (F-26) is

$$x = \exp(At)x^{o} \tag{F-31}$$

(F-30) becomes

$$
\begin{aligned}
H \exp(At)x^o &\equiv 0 \\
HA \exp(At)x^o &\equiv 0 \\
HA^2 \exp(At)x^o &\equiv 0 \qquad \text{(F-32)} \\
\cdots\cdots\cdots \\
HA^{n-1} \exp(At)x^o &\equiv 0
\end{aligned}
$$

If (F-29) is true then (F-32) must be valid for $0 \leq t \leq \infty$. At $t = 0$, then,

$$
\begin{aligned}
Hx^o &= 0 \\
HAx^o &= 0 \\
HA^2x^o &= 0 \qquad \text{(F-33)} \\
\cdots\cdots\cdots \\
HA^{n-1}x^o &= 0
\end{aligned}
$$

(Note: It is now apparent that there is no need to check derivatives of x higher than $(d^{n-1}x)/(dt^{n-1})$, for by the Cayley-Hamilton Theorem A^k $(k \geq n)$ can be found as a linear combination of A^j $(j = 0, 1, 2, \cdots, n-1)$, and if x, $(dx)/(dt)$, $\cdots$, $(d^{n-1}x)/(dt)$ are identically zero, $(d^kx)/(dt^k)$, $(k \geq n)$ must also be identically zero.)

The equations in (F-33) can all be satisfied by a vector $x^o \neq 0$ if and only if

$$
\text{rank} \begin{bmatrix} H \\ HA \\ HA^2 \\ \cdots \\ HA^{n-1} \end{bmatrix} < n \qquad \text{(F-34)}
$$

Thus if

$$\operatorname{rank}\left[H^*,\ A^*H^*,\ (A^*)^2H^*,\ \cdots,\ (A^*)^{n-1}H^*\right] = n \qquad \text{(F-35)}$$

the assumption (F-29) is false and the theorem is proven.

APPENDIX G

ULTRAMINIMAX CONTROL

by

I. Gura
Space Systems Division,
Hughes Aircraft Company
under

Contract No. NAS8-11421

Space Systems Division
AEROSPACE GROUP
Hughes Aircraft Company · Culver City, California

APPENDIX G

ULTRAMINIMAX CONTROL

Derivation of the control law that causes $q \cdot x$ to decay exponentially from an initial perturbation.

Theorem I. For the system

$$\dot{x} = Ax + a\psi \ , \qquad x(0) = x^o \qquad \text{(G-1a)}$$

$$\psi = g \cdot x \qquad \text{(G-1b)}$$

the relationship

$$(q \cdot x) = (q \cdot x^o)\ e^{-\mu t} \qquad \text{(G-2)}$$

for arbitrary q and μ holds, if and only if

$$q \cdot a \neq 0 \qquad \text{(G-3a)}$$

$$g = -\frac{(\mu I + A^*)q}{q \cdot a} \qquad \text{(G-3b)}$$

Furthermore, the closed loop system

$$\dot{x} = Ax + a(g \cdot x) = \tilde{A}x \qquad \text{(G-4)}$$

is asymptotically stable if and only if $\left(q \cdot \Gamma(s)a/q \cdot a\right)$ is Hurwitz.

Proof. Using (G-4)

$$\frac{d(q \cdot x)}{dt} = \left\{A^* q + (q \cdot a)g\right\} \cdot x \qquad \text{(G-5)}$$

However, if (G-2) is to hold

$$\frac{d(q \cdot x)}{dt} = - \mu(q \cdot x) \tag{G-6}$$

Combining (G-5) and (G-6) gives

$$\{A^* q + (q \cdot a)g\} \cdot x = - \mu(q \cdot x) \tag{G-7}$$

This can be valid for all x if and only if

$$g = - \frac{(\mu I + A^*)q}{q \cdot a} \tag{G-8}$$

Using this control it is possible that $q \cdot x$ can decay according to (G-2), while other system variables will grow without bound. To avoid this difficulty, the characteristic equation of the closed loop system must be checked for unstable roots before accepting (G-8) as a useful control law. From Appendix C the closed loop characteristic equation for (G-4) is given by

$$\tilde{\Delta}(s) = \Delta(s) - g \cdot \Gamma(s)a \tag{G-9}$$

Applying (G-8) gives

$$\tilde{\Delta}(s) = \Delta(s) + q \cdot (\mu I + A)\Gamma(s)a/q \cdot a \tag{G-10}$$

Now since

$$(sI - A)^{-1} \triangleq \Gamma(s)/\Delta(s) \tag{G-11a}$$

or equivalently,

$$A\Gamma(s) = s\Gamma(s) - \Delta(s)I \tag{G-11b}$$

Then (G-10) becomes

$$\tilde{\Delta}(s) = \Delta(s) + \frac{q \cdot \mu\Gamma(s)a}{q \cdot a} + \frac{q \cdot s\Gamma(s)a}{q \cdot a} - \Delta(s)\frac{q \cdot a}{q \cdot a} \qquad \text{(G-12a)}$$

or

$$\tilde{\Delta}(s) = (s + \mu)\frac{q \cdot \Gamma(s)a}{q \cdot a} \qquad \text{(G-12b)}$$

Thus (G-8) is useless unless $\left(q \cdot \Gamma(s)a\right)/(q \cdot a)$ is Hurwitz.

Theorem II. Choosing

$\psi = g \cdot x$ in (G-1a) so as to minimize

$$\int_0^\infty \left[\mu^2(q \cdot x)^2 + (g \cdot x)^2\right] dt \qquad \text{(G-13)}$$

as $\mu^2 \rightarrow \infty$ with q arbitrary results in a stable closed loop system identical to that requiring

$$q \cdot x = (q \cdot x^o)\, e^{-\mu t} \qquad \text{(G-14)}$$

if $q \cdot \Gamma(s)a$ is Hurwitz (i.e., ultraminimax control is the same as optimal control in the sense of minimizing (G-13) as $\mu^2 \rightarrow \infty$).

Proof. By Appendix C, minimizing (G-13) gives

$$0 = \tilde{\Delta}(s)\,\tilde{\Delta}(-s) = \Delta(s)\,\Delta(-s) + \mu^2\left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right) \qquad \text{(G-15)}$$

Assume $\Delta(s)\,\Delta(-s)$ is polynomial of degree $2n$ and $\left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right)$ is a polynomial of degree $2m$ where

$$m \leq n - 1 \qquad \text{(G-16)}$$

Obviously, as $\mu^2 \to \infty$, 2m roots of (G-15) approach those of

$$0 = \left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right) \tag{G-17}$$

The remaining 2n - 2m roots are determined as follows.

From (G-15)

$$\frac{\Delta(s)\ \Delta(-s)}{\left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right)} = -\mu^2 \tag{G-18}$$

By long division this can be expressed as

$$s^{2n-2m} + \lambda_1 s^{2n-2m-1} + \cdots + \lambda_{2n-2m}$$

$$+ \lambda_{2n-2m+1} s^{-1} + \cdots + \lambda_{2n} s^{-2m} = -\mu^2 \tag{G-19}$$

where the λ's are constants. Dividing through by s^{2n-2m} and using the complex variable notation

$$-\mu^2 = e^k e^{i\pi} \tag{G-20}$$

(G-19) becomes

$$1 + \lambda_1 s^{-1} + \cdots + \lambda_{2n} s^{-2n} = \frac{e^k e^{i\pi}}{s^{2n-2m}} \tag{G-21}$$

Now if

$$s = (e^k e^{i\pi})^{1/2n-2m} \tag{G-22}$$

(G-21) is satisfied as $e^k \to \infty$. Thus 2n-2m roots of (G-15) are given by the solutions of

$$s^{2n-2m} + \mu^2 = 0 \qquad \text{(G-23}$$

In general then, the roots of (G-15) are given by the roots of

$$0 = \left(s^{2n-2m} + \mu^2\right)\left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right) \tag{G-24}$$

If $q \cdot \Gamma(s)a$ is Hurwitz and $m = n - 1$

$$0 = \tilde{\Delta}(s) = (s + \mu)\left(q \cdot \Gamma(s)a\right) \tag{G-25}$$

is the closed loop characteristic equation of the stable optimal system. This agrees exactly with the ultraminimax system for which

$$q \cdot x = (q \cdot x^o)\, e^{-\mu t} \tag{G-26}$$

Remark. The question arises as to what adjustments can be made if $q \cdot \Gamma(s)a$ fails to be Hurwitz. Consider the equation

$$\left(q \cdot \Gamma(s)a\right)\left(q \cdot \Gamma(-s)a\right) = 0 \tag{G-27}$$

It is clear that m roots of this equation must lie in the left half of the complex plane. From these roots generate the Hurwitz polynomial

$$\sum_{i=0}^{h} \bar{\alpha}_i \cdot s^{i-1} \tag{G-28}$$

where $\bar{\alpha}_m = 1$, and $\bar{\alpha}_i = 0$ for $i > m$. In general, (G-27) will not coincide identically with $q \cdot \Gamma(s)a$.

Now let

$$\alpha_i = \bar{q} \cdot S_i a \tag{G-29}$$

where $\bar{q}$ can be determined from

$$\bar{q} = \sum_{i=1}^{n} \bar{\alpha}_i (A^*)^{i-1} b \tag{G-30}$$

This relationship follows from the identity

$$\left(b,\ A^* b,\ \cdots,\ (A^*)^{n-1} b\right)(S_1 a,\ S_2 a,\ \cdots,\ S_n a)^* = I \tag{G-31}$$

which implies

$$\bar{q} = I\bar{q} = \sum_{i=1}^{n} (\bar{q} \cdot S_i a)(A^*)^{i-1} b \tag{G-32}$$

for any vector $\bar{q}$. From (G-28) then,

$$\sum_{i=0}^{n} \bar{\alpha}_i s^{i-1} \equiv \sum_{i=0}^{n} (\bar{q} \cdot s_i a) S^{i-1} \triangleq \bar{q} \cdot \Gamma(s) a \tag{G-33}$$

Thus when $q \cdot \Gamma(s)a$ is not Hurwitz, there exists a vector $\bar{q}$ such that the polynomial $\bar{q} \cdot \Gamma(s)a$ is Hurwitz. Furthermore, by the results derived in Theorem II, the characteristic equation for a closed loop system minimizing the integral

$$\int_0^{\infty} \left[(q \cdot x)^2 + (g \cdot x)^2\right] dt \tag{G-34}$$

is

$$(s + \mu)\left(\bar{q} \cdot \Gamma(s)a\right) = 0 \tag{G-35}$$

This guarantees that

$$\bar{q} \cdot x = (\bar{q} \cdot x^o)\ e^{-\mu t} \tag{G-36}$$

for all x^o. Note that the system for which (G-35) holds will also minimize (G-13). In a least squares sense, then, $\bar{q} \cdot x$ is the "closest" approximation to $q \cdot x$ which can decay exponentially in a stable closed loop system.

APPENDIX H

CANONICAL FORMS FOR CONTROLLABLE SYSTEMS WITH APPLICATIONS TO OPTIMAL NONLINEAR FEEDBACK

by

R. W. Bass and I. Gura
Space Systems Division
Hughes Aircraft Company

under

Contract No. NAS8-11421

Space Systems Division

AEROSPACE GROUP

Hughes Aircraft Company . Culver City, California

ABSTRACT

FOR

CANONICAL FORMS FOR CONTROLLABLE SYSTEMS WITH APPLICATIONS TO OPTIMAL NONLINEAR FEEDBACK

by

R. W. Bass and I. Gura

Using the assumption of controllability, explicit closed form transformations among four linear canonical forms useful in control system analyses are derived. The relationships found can be easily programmed for efficient numerical computation and are also helpful in obtaining further theoretical results. Indeed, these formula are basic in establishing the properties of a nonlinear canonical form for bang-bang systems, which on each side of the switching surface rectifies the state-space phase portrait of the given system into parallel straight lines. This transformation, in turn, permits direct integration of the Hamilton-Jacobi partial differential equation. Furthermore, the feedback law for the classical time-optimal control problem is shown to have the form of an infinite series of fractional powers of the nonlinear canonical variables.

CANONICAL FORMS FOR CONTROLLABLE SYSTEMS WITH APPLICATIONS TO OPTIMAL NONLINEAR FEEDBACK

by

R.W. Bass
University of Colorado and Hughes Aircraft Company
Culver City, California, U.S.A.
and
I. Gura
Hughes Aircraft Company
Culver City, California, U.S.A.

Notational Conventions

a. Matrices are upper case Roman letters.
b. Vectors are lower case unsubscripted or superscripted letters.
c. Scalars are Greek letters and all subscripted lower case letters.
d. Exceptions to these rules are i, j, k, ℓ, ν, n which are used as summation indices or scalars; $\Gamma(s)$ which is a matrix polynomial; s which is a complex scalar; $\Delta(s)$ which is a polynomial in s; t which denotes time; and σ, ξ, θ, φ which are vectors.
e. Asterisks (*) denote matrix transposition.
f. The i^{th} column of the identity matrix is represented by e^i.
g. The symbol $\triangleq$ denotes equality by definition.
h. Unless otherwise stated, indices will range over the set 1, 2, . . . , n.

Introduction

In the analysis and design of control systems for autonomous linear plants, the utility of simple explicit transformations between the given state variables and certain canonical forms is well known.

It has been shown by Lur'e [1], Letov [2], and many others, that use of Lur'e coordinates facilitates explicit construction of Liapunov functions [3], thus advancing the study of stability of equilibrium in dynamical systems.

More recently it has been shown by Bass, Lewis and Mendelson [4], [5], by Wonham and Johnson [6], [7], [8], by Kalman [9], and by Bass, Gura and Webber [10], [11] that use of phase coordinates permits the fruitful application of frequency-domain concepts to various problems of system stabilization and optimization originally stated in terms of time-domain concepts.

In this paper a system of generalized Lur'e coordinates is introduced. Unlike the Lur'e coordinates, these variables are well-defined even if the system's "open-loop poles" (i.e., the plant's eigenvalues or characteristic roots) are not distinct. Although many realistic engineering problems do not have multiple roots, numerous highly illuminating examples of modern control theory can be derived readily when such roots are permitted. Therefore, the complete generality of applicability of this last-mentioned coordinate system is important for both exposition and research on advanced control problems.

It will be demonstrated below that in both theoretical research and practical design procedures it is rewarding to be able to pass freely between the above-mentioned coordinate systems and the state-space of the given problem. Twelve different linear transformations are needed. Unfortunately, certain key inverse transformations have not been available hitherto in explicit closed form. It has been assumed in previous control work that matrices involved are to be inverted numerically, and thus the needed coefficients were then only defined implicitly. This has led to awkward circumlocutions (e.g., [6], [9])and the desirability of closed form algebraic expressions for the inverses has been widely recognized. Attempts [8], [12] at deriving such expressions in the past have involved unnecessary assumptions (e.g., distinct eigenvalues) and their practical use would (unnecessarily) require computation of both eigenvalues and eigenvectors; in addition, these results have no theoretical utility. Partial objectives of this work are to

1. permit most efficient numerical evaluation of the desired inverses; and to

2. yield theoretical results and new algebraic identities which have facilitated solution of control problems that hitherto appeared formidable, if not intractable.

In Part I below, closed form expressions for all transformations are displayed in systematic arrays. These formulae have been programmed for digital computation and used in the design of an advanced attitude stabilization system for non-rigid aeroballistic vehicles which were "flown" successfully in computer-simulations [13].

Furthermore, some of the new algebraic identities established in Part I have been used in proving various new theoretical results([10], [11], [13], [14] where the identities are stated but not proved). For example, use of phase variables in [10] supplies a direct design procedure which is the inverse of the (indirect) root-locus approach.

A new and evidently important nonlinear transformation, together with its explicit inverse, is introduced in Part II by making free use of the linear canonical forms. This transformation renders trivial the integration of the Hamilton-Jacobi equation pertaining to "bang-bang" optimal feedback control. In fact, the state-space phase-portrait on each side of the switching surface is transformed explicitly into a "rectified" flow along parallel straight lines.

The nonlinear canonical form also permits a constructive solution of the celebrated time-optimal feedback regulator problem. It is shown in Part III that the general time optimal switching function embodies three features noted in the low-dimensional special cases previously solved; namely, the solution is an analytic function of fractional powers of the system's first integrals which can be generated on-line by means of logarithmic amplifiers.

Applications to minimization of quartic and higher order performance indices are also considered.

The system to be studied in this paper is of the type

$$\dot{x} = Ax + a\psi \qquad (1)$$

where

$$\dot{x} = Ax \qquad (2)$$

governs the evolution in time of the uncontrolled plant, where the letter a denotes the actuator vector, and where the scalar function $\psi = \psi(x)$ denotes the feedback control law. In general, the solution of the system of differential equations (1) involves the transition matrix e^{At}, whose Laplace transform is the resolvent matrix $(sI - A)^{-1}$ where I is the identity matrix and s is a complex scalar. It can be shown [4], [15] that this matrix is given by

$$(sI - A)^{-1} = \frac{\Gamma(s)}{\Delta(s)} \qquad (3)$$

where

$$\Delta(s) = \det(sI - A) = \sum_{j=0}^{n} \alpha_j s^j , \quad \Gamma(s) = \sum_{i=1}^{n} s^{i-1} S_i , \qquad (4)$$

and the $S_1, S_2, \ldots, S_n$ and the $\alpha_0, \alpha_1, \ldots, \alpha_n$ are effectively computable by the recursion relations

$$\alpha_n = 1, \quad S_n = I , \qquad (5a)$$

$$\alpha_{n-j} = -\frac{1}{j} \operatorname{trace}(AS_{n-j+1}), \quad S_{n-j} = \alpha_{n-j} I + AS_{n-j+1} . \qquad (5b)$$

The matrices S_i can be shown [4] to satisfy

$$S_{n-j} = \sum_{i=n-j}^{n} \alpha_i A^{i-n+j} . \qquad (5c)$$

The controllability criterion of Kalman [9] is fundamental to the present analysis and will be assumed henceforth. For the system (1) it can be expressed in determinantal form as

$$\det D \neq 0 , \quad D = (a, Aa, \ldots, A^{n-1}a) . \qquad (6)$$

Certainly, if (1) is controllable, the system of simultaneous linear equations

$$a \cdot b = 0, \quad Aa \cdot b = 0 , \ldots , \quad A^{j-1}a \cdot b = 0, \ldots ,$$
$$A^{n-2}a \cdot b = 0, \quad A^{n-1}a \cdot b = 1 \qquad (7)$$

must have a unique vector $b \neq 0$ for its solution. The vector b can be computed by Gaussian elimination. In general, computing b represents $(1/n)^{th}$ of the arithmetic labor required to invert an $n \times n$ matrix. The key inverse matrix desired has columns $(A^*)^{i-1}b$; elementary recursive formulae then supply the other matrices directly.

The vector b is quite remarkable for several reasons. In addition to supplying all five canonical forms presented here, it is fundamentally related to the magnitude of the linear feedback signals required to force (1) to behave in any arbitrary manner [10].

Furthermore, the vector b is the normal vector at x = 0 to the time-optimal switching surface of the given control problem. In fact, it will be proved that near x = 0 the time-optimal regulator law has the form

$$\psi = - \operatorname{sgn} [b \cdot x + \rho(x)] \qquad (8)$$

where $[\rho(x)/ \| x \|] \to 0$ as $\| x \| \to 0$; thus at $x = 0$ the surface $b \cdot x = 0$ is the tangent hyperplane of the surface $b \cdot x + \rho(x) = 0$.

Part I. Linear Canonical Forms

In this section there will be established a complete set of explicit transformations among the canonical forms

Given state variables:	$\dot{x} = Ax + a\psi$,	(9a)
Phase variables:	$\dot{\theta} = C\theta + e^n\psi$,	(9b)
Generalized Lur'e variables:	$\dot{\varphi} = C^*\varphi + e^1\psi$,	(9c)
Lur'e variables:	$\dot{\xi} = \Lambda\xi + u\psi$,	(9d)

where case (9d) is void unless the characteristic roots λ_i of A are distinct, where C is the companion matrix to A, namely

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-2} & -\alpha_{n-1} \end{bmatrix}, \qquad (10)$$

and where

$$\Lambda = (\lambda_1 e^1, \lambda_2 e^2, \ldots, \lambda_n e^n) , \quad u = (1, 1, \ldots, 1)^* . \qquad (11)$$

The forms (9a)-(9c) are real. Since the λ_i occur in complex conjugate pairs, it will be shown that the ξ_i do also; it is easy to put (9d) into a real form in which the complex diagonal matrix Λ is replaced by a real matrix which has 2×2 submatrices along the main diagonal and in which each component of u is either 1 or 0.

Using symbols to be defined as the outline of the derivation proceeds, the desired transformations are as follows.

Coordinate Transformations in Vector-Matrix Form

	x	θ	φ	ξ $(\lambda_i \neq \lambda_j)$
x	$x = x$	$\theta = L^*x$	$\varphi = TL^*x$	$\xi = V^*x$
θ	$x = DT\theta$	$\theta = \theta$	$\varphi = T\theta$	$\xi = Z^*T\theta$
φ	$x = D\varphi$	$\theta = T^{-1}\varphi$	$\varphi = \varphi$	$\xi = Z^*\varphi$
ξ $(\lambda_i \neq \lambda_j)$	$x = DW\xi$	$\theta = T^{-1}W\xi$	$\varphi = W\xi$	$\xi = \xi$

Coordinate Transformations in Vector-Scalar Form

$x = \sum_{i=1}^{n} \theta_i S_i a = \sum_{i=1}^{n} \varphi_i A^{i-1} a = \sum_{i=1}^{n} \xi_i u^i$
$\theta_i = (A^*)^{i-1} b \cdot x = \sum_{j=0}^{i-1} \theta_j \varphi_{j+n-i+1} = \sum_{j=1}^{n} \frac{(\lambda_j)^{i-1}}{\Delta'(\lambda_j)} \xi_j$
$\varphi_i = (S_i^* b) \cdot x = \sum_{j=i}^{n} \alpha_j \theta_{j-i+1} = \sum_{j=1}^{n} \left\{ \sum_{\nu=i}^{n} \frac{(\lambda_j)^{\nu-i}}{\Delta'(\lambda_j)} \alpha_\nu \right\} \xi_i$
$\xi_i = v^i \cdot x = \sum_{j=1}^{n} \left\{ \sum_{k=j} \alpha_k \lambda_i^{k-j} \right\} \theta_j = \sum_{j=1}^{n} \lambda_i^{j-1} \varphi_j$

All of the transformations depend directly on the basic identities

$$L^{-1} \triangleq \left[(b, A^*b, \ldots, (A^*)^{n-1}b)\right]^{-1} \equiv (S_1a, S_2a, \ldots, S_na)^* , \qquad (12)$$

$$D^{-1} \triangleq \left[(a, Aa, \ldots, A^{n-1}a)\right]^{-1} \equiv (S_1^*b, S_2^*b, \ldots, S_n^*b)^* , \qquad (13)$$

$$(L^{-1})^* \equiv DT \quad , \tag{14}$$

where

$$T = T^* \triangleq \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & 1 \\ \alpha_2 & \alpha_3 & \cdots & 1 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \alpha_{n-1} & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad . \tag{15}$$

The inverse of T is

$$T^{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & \beta_1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 1 & \cdots & \beta_{n-3} & \beta_{n-2} \\ 1 & \beta_1 & \cdots & \beta_{n-2} & \beta_{n-1} \end{bmatrix} \tag{16}$$

where the β_i are defined by the Laurent series

$$\frac{1}{\Delta(s)} \equiv \sum_{j=0}^{\infty} \frac{\beta_j}{s^{n+j}} \, , \qquad (|s| > \max|\lambda_i|) \, , \tag{17}$$

and can be calculated by the recursion relations

$$\beta_0 = 1, \quad \beta_\ell = -\sum_{j=0}^{\ell-1} \alpha_{j+n-\ell}\,\beta_j, \quad (\ell = 1, 2, \ldots, n) \tag{18a}$$

$$\beta_{\ell+n} = -\sum_{j=\ell}^{\ell+n-1} \alpha_{j-\ell}\,\beta_j \, , \quad (\ell = 1, 2, 3, \ldots) \, . \tag{18b}$$

To prove (12), consider the equivalent form $e^i = (S_1 a, S_2 a, \ldots, S_n a)^* (A^*)^{i-1} b$, which, under row by row expansion and application of (5c) can be expressed as

$$a^* \sum_{k=i}^{n+i-j} \alpha_{k+j-i} (A^*)^{k-1} b = \delta_{ij} \quad . \tag{19}$$

Now from (7) expressed in the form

$$\delta_{kn} = a^* (A^*)^{k-1} b \quad , \tag{20}$$

it is easy to show that (19) is valid for $1 \le k \le n$, or when $j \ge i$. When $j < i$ expand the left side of (19) into two parts, the first consisting of the terms for which $i \le k \le n$, and the second consisting of the remaining terms $(n < k \le n + i - j)$. Then by using (20) in the first part, and the Cayley-Hamilton Theorem in the second, the proof can be completed.

The identity (14) follows directly from explicit expansion of the matrix product DT and application of (5c) and (12). Similarly, identity (13) comes from the expansion of LT and the use of (5c) and (14).

The relationships (18) can be verified by manipulating (17) into the form

$$1 \equiv \sum_{\nu=0}^{\infty} \left\{ \sum_{i=\max(n-\nu,\,0)}^{n} \alpha_i \beta_{i+\nu-n} \right\} s^{-\nu} \tag{21}$$

and then comparing coefficients of $s^{-\nu}$ on both sides of the equation. Using (18a), (16) can be established by direct matrix multiplication of T and T^{-1}.

A. Phase Variables (θ)

Consider the output of interest for the system (1) to be

$$\theta_1 \triangleq b \cdot x \quad . \tag{22}$$

By alternately differentiating (22) and applying (1) and (7), the relationships

$$\frac{d^{i-1}\theta_1}{dt^{i-1}} = (A^*)^{i-1} b \cdot x \quad , \qquad \frac{d^n \theta_1}{dt^n} = (A^*)^n b \cdot x + \psi \tag{23}$$

can be established. Then, multiplying the j^{th} derivative of θ_1 by α_j, summing over $j = 0, 1, 2, \ldots, n$, and applying the Cayley-Hamilton Theorem gives

$$\sum_{j=0}^{n} \alpha_j \frac{d^j \theta_1}{dt^j} = \Delta(d/dt)\theta_1 = \psi \quad . \tag{24}$$

Upon defining the state variables $\theta_1, \theta_2, \ldots, \theta_n$ by

$$\theta_i = d^{i-1}\theta_1 / dt^{i-1} \quad , \tag{25}$$

the n^{th} order scalar differential equation (24) can be expressed as the first order matrix system (9b). On combining (23) with (25), it is obvious that

$$\theta = [b, A^* b, \ldots, (A^*)^{n-1} b]\, x = L^* x \quad . \tag{26}$$

Note that applying this directly to (1) and comparing the result with (9b) shows that $C = LA(L^*)^{-1}$. The identities of (12), (13), and (14) can be used to give the explicit inverse of (26), namely

$$x = (L^*)^{-1}\theta = (S_1 a, S_2 a, \ldots, S_n a)\theta = \sum_{i=1}^{n} \theta_i S_i a, \quad x = DT\theta \, . \tag{27}$$

B. Generalized Lur'e Variables (φ)

Let

$$\varphi \triangleq D^{-1} x = TL^* x \quad . \tag{28}$$

Then (1) becomes

$$\dot\varphi = (D^{-1} A D)\varphi + D^{-1} a\psi \quad . \tag{29}$$

Upon forming the product DC^*, and applying the Cayley-Hamilton Theorem to the result, it becomes obvious that $C^* = D_1^{-1} AD$. From (13), it can be shown that $D^{-1} a = e^1$, whence (29) is equivalent to (9c). Combining (28) and (13) gives the inverse transformations

$$\varphi_i = (S_i^* b) \cdot x \quad , \tag{30}$$

$$x = D\varphi = (a, Aa, \ldots, A^{n-1} a)\varphi = \sum_{i=1}^{n} \varphi_i A^{i-1} a \quad .$$

To find the relationship between φ and θ apply (27) to (28), obtaining

$$\varphi = T\theta \quad , \quad \theta = T^{-1}\varphi \quad . \tag{31}$$

The corresponding vector-scalar formulas shown in the table above can be derived directly from these relations and the basic properties of the α_i and β_i. The details are somewhat involved but quite straightforward.

C. Lur'e Variables (ξ)

Consider the φ coordinates for a system with distinct complex eigenvalues λ_i, ($i = 1, 2, \ldots, n$). Multiplying both sides of the vector system (9c) by the transpose of the Vandermonde Matrix $Z = (z^1, z^2, \ldots, z^n)$, where

$$z^i = \sum_{k=1}^{n} (\lambda_i)^{k-1} e^k \quad , \tag{32}$$

and simplifying the result by using $\Delta(\lambda_i) = 0$ yields

$$\sum_{j=1}^{n} \lambda_i^{j-1}\dot{\varphi}_j = \sum_{j=1}^{n} \lambda_i^j \varphi_j + \psi \quad . \tag{33}$$

Defining a new vector ξ by

$$\xi \triangleq Z^*\varphi \tag{34}$$

or equivalently,

$$\xi_i \triangleq \sum_{j=1}^{n} \lambda_i^{j-1}\varphi_j \quad , \tag{35}$$

(33) yields (9d). The inverse of the matrix Z^* can be shown to be $W = (w^1, w^2, \ldots w^n)$ where

$$w^i = Tz^i/\Delta'(\lambda_i) = \sum_{j=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \sum_{k=j}^{n} \alpha_k e^{k-j+1} \quad , \tag{36}$$

with $\Delta'(\lambda_i)$ denoting the polynomial $d(\Delta(s))/ds$ evaluated at $s=\lambda_i$. Clearly then

$$\varphi = W\xi, \quad W = (Z^*)^{-1} \quad . \tag{37}$$

The relationships between ξ and θ, namely

$$\theta = T^{-1}W\xi \quad , \tag{38a}$$

$$\xi = Z^*T\theta \tag{38b}$$

follow from (37), (31), and (34). Details of the development of the corresponding vector-scalar forms are omitted.

Combining (30) and (35) the relationship between x and ξ is seen to be

$$\xi = V^*x \quad , \quad \xi_i = v^i \cdot x \quad , \tag{39}$$

where

$$V = (v^1, v^2, \ldots, v^n) \ , \quad v^i = \sum_{j=1}^{n} \lambda_i^{j-1} S_j^* b = \Gamma^*(\lambda_i)b. \tag{40}$$

Alternatively, from (14), (30), and (34), $\xi = Z^*TL^*x$, so that

$$V = Z^*TL^* \ . \tag{41}$$

By (14) and (37), the inverse relationship is

$$x = U\xi \quad , \quad U \triangleq (Z^*TL^*)^{-1} = DW \quad . \tag{42}$$

Expansion shows the i^{th} column of U to be

$$u^i = \sum_{j=1}^{n} \left\{ \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \right\} S_j a = \frac{\Gamma(\lambda_i)}{\Delta'(\lambda_i)} a \quad . \tag{43}$$

D. An Alternative Generalization of the Lur'e Variables ($\xi(s)$)

The identity

$$\Delta(\eta) - \Delta(\mu) \equiv (\eta - \mu) \sum_{i=1}^{n} \eta^{i-1} \sum_{j=i}^{n} \alpha_j \mu^{j-i} \tag{44}$$

can easily be verified by equating coefficients of like powers of η and μ where these quantities obey the commutative and distributive laws of algebra. With no loss of generality, η can be identified with sI and μ with the matrix A. Then, by the Cayley-Hamilton Theorem and the definition of $\Gamma(s)$,

$$\Delta(s)I = (sI - A)\Gamma(s) \quad . \tag{45}$$

Indeed, (3) can be found directly from this relationship whenever $(sI - A)^{-1}$ exists. Multiplying (45) on the right by the vector a, applying the definition

$$u(s) \triangleq \frac{\Gamma(s)a}{\Delta(s)} \quad , \tag{46}$$

and using (4), it can be seen that, identically in s,

$$Au(s) = s\,u(s) - a \quad , \tag{47a}$$

$$u(s) \cdot b = \frac{1}{\Delta(s)} \quad . \tag{47b}$$

Similarly, considering (44) again with A^* as μ, and using

$$v(s) \triangleq \Gamma^*(s)b \quad , \tag{48}$$

the identities

$$A^*v(s) = s\,v(s) - \Delta(s)b \quad , \tag{49a}$$

$$v(s) \cdot a = 1 \tag{49b}$$

can be derived.

Now define

$$u^i \triangleq \lim_{\rho \to 0} \frac{1}{2\pi\sqrt{-1}} \oint_{|s-\lambda_i|=\rho} u(s)\,ds \quad , \tag{50a}$$

$$v^i \triangleq v(\lambda_i) = \Gamma(\lambda_i)\,b \quad , \tag{50b}$$

and note that when the λ_i are distinct,

$$u^i = \frac{\Gamma(\lambda_i)}{\Delta'(\lambda_i)} a \quad . \tag{50c}$$

Applying the contour integral operator of (50a) to (47a, b), and inserting $s = \lambda_i$ in (49a, b), one obtains, for the case of distinct λ_i,

$$Au^i = \lambda_i u^i \quad , \quad u^i \cdot b = 1/\Delta'(\lambda_i) \quad , \tag{51a}$$

$$A^*v^i = \lambda_i v^i \quad , \quad v^i \cdot a = 1 \quad . \tag{51b}$$

Furthermore, comparing (50c) with (43), and (48) with (40), it becomes clear that the columns of U are the eigenvectors of A normalized by the scaling requirement $u^i \cdot b = 1/\Delta'(\lambda_i)$, and that the columns of V are the eigenvectors of A^* normalized by the scaling requirement $v^i \cdot a = 1$. Since standard digital computer routines do not normalize the lengths of the eigenvectors u^i and v^i in this manner, care must be taken to multiply

u^i by $[1/(u^i \cdot b)\Delta'(\lambda_i)]$, and to multiply v^i by $[1/(v^i \cdot a)]$ (which is permissible since neither denominator vanishes, by the hypotheses of controllability and distinct roots). This discovery that the Lur'e canonical form is precisely equivalent to the standard diagonalization procedure when normalized as in (51a, b) is practically useful in numerical work.

Note that (39) can now be generalized, using (48) and (30), to

$$\xi(s) = v(s) \cdot x = \sum_{i=1}^{n} s^{i-1} \varphi_i \quad . \tag{52}$$

Then, taking the scalar product of v(s) with the system (1) and applying (49a, b), it is found that

$$v(s) \cdot \dot{x} = x^*(sv(s) - \Delta(s)b) + \psi \quad . \tag{53}$$

Now using (52) and (23), the above becomes

$$\dot{\xi}(s) = s\xi(s) - \Delta(s)\theta_1 + \psi \ , \qquad \theta_1 = b \cdot x = \varphi_n \quad . \tag{54}$$

In Part II the system (1) will be considered in the form (54), which is equivalent to (9c) and may be regarded as another generalization of the Lur'e canonical form. In fact, when the eigenvalues of A are distinct, $\xi_i = \xi(\lambda_i)$, and, setting $s = \lambda_i$ in (54), the Lur'e system (9d) is recovered. On the other hand, whether or not the λ_i are distinct, the identity (54), which, in appearance, is highly reminiscent of the Lur'e form, can be regarded as the collection of n differential equations obtained by equating like powers of s on the right and left hand sides. Then, on inserting (52) into (54) and comparing coefficients, the canonical form (9c) can be recovered immediately. It is for this reason that (9c) was called the "Generalized Lur'e Canonical Form."

Part II. A Nonlinear Canonical Form

In this section it will be shown that the (real) systems

$$\dot{x} = Ax + a\epsilon \ , \qquad (\epsilon = \pm 1) \ , \tag{55a}$$

and

$$\dot{\sigma} = \epsilon e^n \ , \qquad (\epsilon = \pm 1) \ , \tag{55b}$$

are related near $x = \sigma = 0$ by the uniquely reciprocal (real) transformations

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} \frac{\epsilon}{s} \log\left[1 + \epsilon s \xi(s)\right] ds \ , \tag{56a}$$

$$\xi(s) = v(s) \cdot x \ ,$$

$$x = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{\epsilon}{s} \left\{ \exp\left[\epsilon \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} s^{\ell+1} \sigma \right] - 1 \right\} u(s)\, ds, \tag{56b}$$

where the path of integration is a circle enclosing all the roots of $\Delta(s) = \Delta(\lambda_i) = 0$; that is, $\rho > \max|\lambda_i|$. For systems with distinct eigenvalues, (56a) and (56b) become, respectively

$$\sigma_j = \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \left\{ \frac{\epsilon}{\lambda_i} \log\left[1 + \epsilon \lambda_i \xi_i\right] \right\}, \quad \xi_i = v^i \cdot x \ , \tag{57a}$$

$$x = \sum_{i=1}^{n} \frac{\epsilon}{\lambda_i} \left\{ \exp\left[\epsilon \sum_{\nu=1}^{n} \sum_{\ell=0}^{n-\nu} \alpha_{\nu+\ell} \lambda_i^{\ell+1} \sigma_\nu \right] - 1 \right\} u^i \ . \tag{57b}$$

Furthermore, it will be shown that the transformation (56a) can always be expressed by an eigenfunction expansion

$$\sigma_i = \sum_{\ell=0}^{\infty} \beta_\ell \omega_{\ell+n-i+1}(\varphi) \ , \qquad \varphi_j = (S_j^* b) \cdot x \ , \tag{58}$$

where the β_ℓ satisfy (18), and the eigenfunctions $\omega_\nu = \omega_\nu(\varphi)$ are multinomials of degree ν in $\varphi_1, \varphi_2, \ldots, \varphi_n$, also recursively computable by

$$\omega_1 = \varphi_1 \ , \quad \omega_\nu = \varphi_\nu - \frac{\epsilon}{\nu} \sum_{m=1}^{\nu-1} m\, \omega_m \varphi_{\nu-m} \ , \quad (\nu = 2, \ldots, n), \tag{59a}$$

$$\omega_{\nu+n} = -\frac{\epsilon}{\nu+n} \sum_{i=1}^{n} (\nu+n-i)\varphi_i \omega_{\nu+n-i} \ , \quad (\nu = 1, 2, 3, \ldots). \tag{59b}$$

Note that $\beta_\ell = \beta_\ell(A)$, and $\varphi_j = \varphi_j(x) = \varphi_j(x;\ A,\ a)$ but that the multinomials $\omega_\nu = \omega_\nu(\varphi)$ depend only on the dimension n of the system and therefore can be computed and tabulated once and for all.

The transformation (56a) also can be expressed by a power series expansion

$$\sigma_j = (A^*)^{j-1}\, b \cdot x - \frac{1}{2}\epsilon\,(x \cdot Q_j x) + \ldots \ , \tag{60a}$$

where

$$Q_1 \triangleq (D^{-1})^* \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 & \beta_1 \\ 0 & 0 & \cdots & 1 & \beta_1 & \beta_2 \\ 0 & 0 & \cdots & \beta_1 & \beta_2 & \beta_3 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 1 & \beta_1 & \cdots & \beta_{n-3} & \beta_{n-2} & \beta_{n-1} \\ \beta_1 & \beta_2 & \cdots & \beta_{n-2} & \beta_{n-1} & \beta_n \end{bmatrix} D^{-1} \tag{60b}$$

$$Q_{j+1} \triangleq A^* Q_j \ , \qquad (j = 1, 2, \ldots, n-1) \quad . \tag{60c}$$

In the case of distinct roots λ_i, alternative expressions for the power-series coefficients are

$$(A^*)^{j-1} b = \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} v^i \ , \tag{60d}$$

$$Q_j = (A^*)^{j-1} Q_1 = \sum_{i=1}^{n} \frac{(\lambda_i)^j}{\Delta'(\lambda_i)} v^i (v^i)^* \quad . \tag{60e}$$

The authors have simulated approximately time-optimal systems of order n = 2, 3, 4, 5 on analog computers by each of the three nonlinear canonical form approaches (57a), (58), and (60a), and have experience in the numerical use of (18), (59a, b), (60b, c), and (60d, e). On combining the complex conjugate terms in (57a), it can be seen that on-line mechanization of σ_j can be effected in an analog control-computer using nonlinear amplifiers which over a suitable dynamic range provide the logarithm, exponential, sine, and cosine. Use of solid-state devices of known nonlinear characteristics (e.g., Zener diodes), or piece-wise linear approximation of the σ_j by biased-diode function generators may prevent $\sigma_j = [(A^*)^{j-1} b] \cdot x + \ldots$ for holding for small $\|x\|$. Therefore (60a) is desirable for small $\|x\|$. However (60a) in the form $\sigma_j \cong [(A^*)^{j-1} b] \cdot x$ does not yield stability in general (unless the vector b is "tilted" to compensate for the absent quadratic and higher terms), nor does even the form $\sigma_j \cong [(A^*)^{j-1}b]\cdot x - \frac{1}{2}\epsilon(x\cdot Q_j x)$ yield asymptotic stability for unstable plants unless Q_1 is modified slightly for similar reasons. The fact that the required modification in Q_1 is less than that needed in b suggests that perhaps extension of (60a) to include the cubic terms in x would constitute a practically adequate (local) mechanization of (57a). The truncation properties of (58) are quite different. Recalling that $\beta_o = 1$, and defining $\rho > \max|\lambda_i|$, it can be shown that as $\rho \to 0$, $\beta_\nu \to 0$ for $\nu = 1, 2, 3, \ldots$. When $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, $\sigma_i \equiv \omega_{n-i+1}(\varphi(x))$ and the truncation of the series (58) at its first term is rigorously valid.

The nature of the preceding transformations depends on the theory of "integrals" and "isochrones." A first integral of the n^{th} order system

$$\dot{x} = f(x) \quad , \qquad x(0) = x^o \tag{61}$$

is a scalar function $\zeta(x)$ such that

$$\zeta[x(t)] \equiv \zeta(x^o) \quad , \qquad \text{or} \qquad f(x)\cdot\operatorname{grad}\zeta(x) \equiv 0 \tag{62}$$

is satisfied along any solution of (61). Geometrically, (62) defines an integral surface such that any state space trajectory initiating on it must remain on it for all t. The term "integral" is used interchangeably for the function $\zeta(x)$ and the surface $\zeta(x)$ = constant.

An isochrone is a surface defined by setting the scalar function $\gamma(x)$ = constant where $\gamma(x)$ satisfies

$$\gamma[x(t)] \equiv \gamma(x^o) + t \quad , \text{or} \quad f(x)\cdot\operatorname{grad}\gamma(x) \equiv 1 \quad . \tag{63}$$

The time for points on various trajectories to move between fixed isochrones is constant; hence the term "isochrone." Here also, this term can refer to either the function $\gamma(x)$ or the surface $\gamma(x) \equiv$ constant.

The following basic properties of integrals and isochrones are readily proved.

1. Any arbitrary function of integrals is also an integral.
2. Every integral of an n^{th} order system can be expressed in terms of any n-1 functionally independent integrals in the neighborhood of a non-equilibrium point. (Proof is analogous to the one of [16], p. 115.)
3. The sum of an integral and an isochrone is an isochrone.
4. Every isochrone of an n^{th} order system can be expressed as the sum of an arbitrary function of n-1 functionally independent integrals and any particular isochrone.

Clearly, the $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ defined by (55b) are first integrals of (55a), while σ_n is an isochrone of that system. It will be shown below that as a consequence of controllability the $\sigma_1, \sigma_2, \ldots, \sigma_n$ are indeed functionally independent so that all of the above properties apply to these functions. The transformations discussed here can be viewed as methods of generating integrals and isochrones for (55a), instead of relationships between canonical variables. This alternative viewpoint is fundamental to analysis of the time-optimal problem.

A. Transformation from x to σ

Differentiate (56a) and apply (54) to obtain

$$\dot{\sigma}_j = \frac{\epsilon}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} \frac{[\epsilon s\xi(s) + \epsilon^2]\, ds}{1 + \epsilon s\xi(s)} - \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}\theta_1\, ds}{1 + \epsilon s\xi(s)} \quad . \tag{64}$$

Now, by complex integration as $\rho \to \infty$, the first term of the right side of (64) becomes

$$\frac{\epsilon}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} = \epsilon\delta_{jn} \quad , \tag{65}$$

while the remaining term can be expressed as

$$\frac{\theta_1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} s^{j-1} \left\{ \sum_{k=0}^{\infty} (-1)^k [\epsilon s\xi(s)]^k \right\} ds \equiv 0 \ , \tag{66}$$

if $|\rho\xi(\rho)|<1$. (Note that the condition $|\rho\xi(\rho)|<1$ can be obtained as a constraint on $||x||$ by applying (52) to obtain $||x||<1/\rho\bar{v}(\rho)$ where $\bar{v}$ is the upper bound of $||v(s)||$ on $|s| = \rho$.) The above result is obvious, since the integrand is analytic in s. Thus (56a) does indeed yield (55b) when applied to (55a).

B. Transformation from σ to x

The simplest proof of (56b) seems to be that in which (57b) is proved first, independently; and then used to establish the more general result.

Consider (57a) and define a vector q such that each component is given by

$$q_i = \frac{\epsilon}{\Delta'(\lambda_i)} \frac{1}{\lambda_i} \log\left[1 + \epsilon\lambda_i\xi_i\right] \quad . \tag{67}$$

Then (57a) can be expressed in vector-matrix form as $\sigma = Zq$, where Z is the Vandermonde Matrix (31). Applying the inverse of Z, (36), $q = W^*\sigma$, or $q^i = w^i\cdot\sigma$, and (67) can be written as

$$\xi_i = \frac{\epsilon}{\lambda_i} \left\{ \exp\left[\epsilon\lambda_i\Delta'(\lambda_i)\, w^i\cdot\sigma\right] - 1 \right\} \quad . \tag{68}$$

By expanding w^i as in (36) and rearranging, (68) becomes the first desired result (57b).

Now define a transformation $x = h(\sigma, \epsilon)$ by (56b). Using the Calculus of Residues, it is clear that (57b) is equivalent when the λ_i are distinct. Also, for distinct λ_i, (56a) and (57a) are equivalent. Hence it is certain that (56b) is the inverse of (56a), at least when the λ_i are distinct. It will now be shown that this proposition is valid for all systems, even when the λ_i are non-distinct. To accomplish this, consider (56a) in the form $\sigma = p(x; \epsilon, A, a)$ and define A to be simple when the roots of its characteristic polynomial $\Delta(s)$ are distinct. It is well-known that if A is not simple there are simple matrices A_o such that $||A-A_o||$ is arbitrarily small. Thus it has been shown that there exists a function $h(\sigma; \epsilon, A, a)$, namely (56b) such that $h(p(x;\epsilon, A, a); \epsilon, A) \stackrel{\Delta}{=} \hat{h}(x;\epsilon, A, a) = x$ is valid whenever A is simple. Now take A non-simple. Let $\{A_\nu\}$ be a sequence such that A_ν is simple for each $\nu = 1, 2, 3, \ldots$ and such that $A_\nu \to A$ as $\nu \to \infty$. Now the integrand in (56a) is a continuous function of x, A, a, and ϵ since $v(s)$ is a polynomial in A, a, and $1/\Delta(s)$. Recall also that $1/\Delta(s)$ is an infinite series in powers of s^{-1}, which converges for $|s|>\max(\lambda_i)$, whose coefficients are rational functions of A. Thus $p(x;\epsilon, A, a)$ is a continuous function of all its arguments. Clearly, an analogous result can be obtained for $h(\sigma;\epsilon, A, a)$. Thus $\hat{h}(x;\epsilon, A, a)$ is continuous in all arguments and so $\hat{h}(x;\epsilon, A_\nu, a) \to \hat{h}(x;\epsilon, A, a)$ as $\nu\to\infty$. But since $\hat{h}(x;\epsilon, A_\nu, a) \equiv x$, it follows upon taking the limit that $\hat{h}(x;\epsilon, A, a) = x$. This completes the proof of the validity of (56b) as the general inverse of (56a).

C. Expansion of σ in Series of Recursively Computable Multinomials

Consider the Taylor expansion

$$\epsilon \log(1+\epsilon s\xi(s)) = \sum_{j=1}^{\infty} \epsilon \frac{(-1)^{j+1}}{j}[\epsilon s\xi(s)]^j \ , \qquad (69)$$

which holds for $|\epsilon s\xi(s)|<1$. Since, by (52), $\xi(s)$ is a polynomial in s, the right side of (69) must be an infinite series in s and so

$$\epsilon \log[1+\epsilon s\xi(s)] = \sum_{j=1}^{\infty} \omega_j s^j \ , \qquad (70)$$

where the coefficients ω_j, $(j = 1, 2, 3, \ldots)$, are to be determined. To accomplish this end, differentiate both sides of (70) with respect to s, apply (52) and simplify, obtaining

$$\sum_{i=1}^{n} i s^{i-1}\varphi_i = \sum_{j=1}^{\infty} j\omega_j s^{j-1} + \epsilon \sum_{k=2}^{\infty} \sum_{i=1}^{\min(k-1,n)} (k-i)\varphi_i\omega_{k-i}s^{k-1} \ . \qquad (71)$$

Then, equating like coefficients in (71), the recursion relations (59) can be established. Now note that, using (70), (56a) can be expressed as

$$\sigma_i = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{i-1}}{\Delta(s)} \frac{1}{s} \sum_{j=1}^{\infty} \omega_j s^j \, ds \ . \qquad (72)$$

From the series expansion of $1/\Delta(s)$ given in (17) and the Calculus of Residues, (72) yields the desired result (58).

Example: the n-Fold Integrator

The system

$$d^n\theta_1/dt^n \equiv \theta_1^{[n]} = \epsilon \qquad (73)$$

was treated by Lewis and Mendelson [5] for n = 3, 4, but no systematic procedure for calculating the integrals and isochrones of (73) was given. By application of (58)—(59), it becomes a simple matter to do so. Since the characteristic equation for (73) is $\Delta(s) = s^n$, $\alpha_0 = \alpha_1 = \ldots = \alpha_{n-1} = 0$. Then from (18) it is clear that $\beta_\ell = 0$, $(\ell = 1, 2, 3, \ldots)$, and so (58) becomes $\sigma_i = \omega_{n-i+1}$. Also, by the vector-scalar relationships between φ and θ, $\varphi_i = \theta_{n-i+1} = \theta^{[n-1]}$, $(i = 1, 2, \ldots, n)$. Thus (59) yields

$$\sigma_n = \theta_1^{[n-1]} \ , \qquad (74a)$$

$$\sigma_i = \theta_1^{[i-1]} - \frac{\epsilon}{n-i+1} \sum_{m=1}^{n-i} m\sigma_{n-m+1}\theta_1^{[i+m-1]} \ ,$$

$$(i = 1, 2, \ldots, n-1) \ . \qquad (74b)$$

D. Power Series Expansion of σ

Expressing the integrand of (56a) in a power series in ξ, and applying the expanded form of (38a) and (23), results in

$$\sigma_j = (A^*)^{j-1} b\cdot x - \frac{1}{2}\epsilon \sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} \lambda_i \xi_i^2 + \ldots \ . \qquad (75)$$

for $|\lambda_i\xi_i|<1$. Define Q_j as in (60e) and apply (39) to (75) to obtain the quadratic terms in the form $-1/2\,\epsilon(x\cdot Q_j x)$. The relationship $A^* Q_j = Q_{j+1}$ directly follows from (60e) and (51b). To obtain a more explicit representation of Q_1, note that on using (41) and (14), (60e) (with j=1) can be expressed in the form

$$Q_1 = (D^{-1})^* E D^{-1} \ , \qquad (76a)$$

$$E \triangleq \sum_{i=1}^{n} \frac{\lambda_i}{\Delta'(\lambda_i)} z^i (z^i)^* \ . \qquad (76b)$$

Then by (32) and the relation

$$\sum_{i=1}^{n} \frac{(\lambda_i)^{j-1}}{\Delta'(\lambda_i)} = \beta_{j-n} \ , \quad (j = 1, 2, 3, \ldots) \ , \qquad (77)$$

(obtained by contour integration of $\oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)} ds$) the $(\nu, \mu)^{th}$ element of E is

$$e^\nu\cdot E e^\mu = \sum_{i=1}^{n} \frac{(\lambda_i)^{\nu+\mu-1}}{\Delta'(\lambda_i)} = \beta_{\nu+\mu-n} \ , \quad (\nu,\mu = 1, \ldots, n). \qquad (78)$$

Thus (60a) is verified for systems with distinct eigenvalues. To generalize the above proof, note that (56a) is analytic in a neighborhood of x = 0, and so there must exist vectors $\ell^i = \ell^i(A, a)$ and matrices $R_i(A, a)$ such that

$$\sigma_j = \ell^j\cdot x - \frac{1}{2}\epsilon(x\cdot R_j x) + \ldots \ , \qquad (79)$$

for all A. Furthermore, $\ell^i(A, a)$ and $R_i(A, a)$ are rational functions of the elements of (A, a). But the expressions in (60b, c) are well-defined rational functions of (A, a) whether or not A is simple, and it has just been proved that

$$\ell^j = (A^*)^{j-1} b \ , \quad R_j = (A^*)^{j-1} Q_1 \ , \qquad (80)$$

whenever A is simple. Hence by the continuity argument used after (68) the relationships (80) must remain valid for all matrices A, simple or not.

Previously, it was claimed that the elements of σ are functionally independent at x = 0. This can be proved by means of the series representation for σ. The Jacobian Matrix for the transformation in question is, by (60a), $L = (b, A^*b, \ldots, (A^*)^{n-1}b$. From (14), however, det L = det D, hence L is not singular if the system (1) is controllable.

Part III. Optimal Nonlinear Feedback Control

Imposing an inequality constraint upon the control function ψ, consider the problem of choosing ψ in

$$\dot{x} = Ax + a\psi \ , \quad |\psi| \le 1 \ , \quad x(0) = x^\circ \ , \qquad (81)$$

so as to minimize a performance criterion

$$\Phi = \Phi(x^\circ) = \int_0^\tau \Psi(x)\,dt \ , \quad (\Psi > 0 \text{ if } x \neq 0) \ , \qquad (82)$$

where the stopping time $\tau = \tau(x^\circ) \le +\infty$ is defined by $x(t) \to 0$ as $0 \le t \to \tau$.

The Hamilton-Jacobi Equation and Liapunov Stability

Assume that an optimal control law $\psi = \psi(x)$ is known, and that $\Phi(x)$ and $\tau(x)$ are continuously differentiable. Obviously (82) is a solution of the partial differential equation

$$\dot{\Phi} \triangleq (Ax + a\psi)\cdot \operatorname{grad} \Phi = -\Psi \ (<0 \text{ if } x \neq 0) \ , \qquad (83)$$

because $d\Phi(x(t))/dt = \dot{\Phi}(x)$ when $x(t)$ satisfies (81). Similarly, if $\Psi(x)$ and $\psi(x)$ are such that (83) has a positive definite solution $\Phi(x) > 0$ if $x \neq 0$, with $\Phi(x) \to +\infty$ as $||x|| \to +\infty$, then either there exists a stopping time τ, or else (if ψ is not everywhere continuous) a time $\tau_\infty(x^\circ)$ such that the solution of (81) cannot be conventionally defined for $t > \tau_\infty$. (For the theory of "chattering" or "after-end-point motion" or the "sliding regime," see [17].) Note that (83) can be expressed as

$$\mathcal{H} = \mathcal{H}(x, y, \psi) = 0 \ , \quad \mathcal{H} \triangleq y\cdot(Ax + a\psi) - \Psi(x) \ , \qquad (84)$$

$$y = -\operatorname{grad} \Phi(x) \ . \qquad (85)$$

Principle of Optimality and the Maximum Principle

If the problem (81)–(82) has a solution then the Maximum Principle, which has been proved rigorously [18], asserts that as a necessary condition there exists, for fixed x°, a function $y = y(t) = y(t;x^\circ)$ such that not only (84) holds, but moreover

$$\mathcal{H} = \overline{\mathcal{H}}(x, y) = 0 \ , \quad \overline{\mathcal{H}} = \max_{|\psi| \le 1} \mathcal{H}(x, y, \psi) \ . \qquad (86)$$

$$\dot{x} = \operatorname{grad}_{(y)} \mathcal{H}(x, y, \psi) \ , \quad \dot{y} = -\operatorname{grad}_{(x)} \mathcal{H}(x, y, \psi) \ . \qquad (87)$$

However, (85) is not claimed to be necessary. Starting from the valid Principle of Optimality [19], a formal, heuristic argument indicates that (84)–(85)–(86) are both necessary and sufficient. But rigorous study of (84)–(87) is difficult. Clearly (86) implies that

$$\psi = \operatorname{sgn}[\sigma_0(x)] \ , \quad \sigma_0(x) \triangleq -a\cdot \operatorname{grad} \Phi \neq 0 \ , \qquad (88)$$

whence there is a hypersurface, $\sigma_0 = 0$, along which $\psi(x)$ is discontinuous; on either side of this surface, ψ is a constant. It is easy to prove that if Φ satisfies (84)–(85)–(86) in the complement of the set $\sigma_0(x) = 0$, then the known necessary condition (87) is a corollary. However, the definition of ψ on the set $\sigma_0 = 0$ is difficult, as is the extension of the just mentioned result about (87) onto the set $\sigma_0 = 0$. In some problems ψ must be given the value $+1$ or -1 on various portions of the set $\sigma_0 = 0$, so that $\sigma_0 = 0$ constitutes an integral surface. Other problems [11] allow two equally valid alternatives: (i) ψ can be defined as a continuous function such that $\sigma_0 = 0$ is an integral surface; or (ii) ψ can be regarded as zero on $\sigma_0 = 0$, and yet the "chattering regime" governed by (88) yields an $x(t)$ identical to that of (i). This phenomenon is connected with the singular solutions of (81)–(82), along which $a\cdot y(t) \equiv 0$, and singular surfaces of (84)–(86) on which $a\cdot \operatorname{grad} \Phi(x) = 0$. Choosing alternative (ii) unifies the two kinds of problems under the subject of bang-bang control, wherein

$$\dot{x} = Ax + a\epsilon \ , \quad \epsilon = \epsilon(x) \ , \quad \epsilon^2 \equiv 1 \ . \qquad (89)$$

Denoting (56b) by $x = h(\sigma)$, and defining $\hat{\Phi}(\sigma) \triangleq \Phi(h(\sigma))$ and $\hat{\Psi}(\sigma) \triangleq \Psi(h(\sigma))$, the system (89) becomes $\dot{\sigma} = \epsilon e^n$ and the equations (84), (85) and (88) become

$$(\partial\hat{\Phi}/\partial\sigma_n) = -\epsilon\hat{\Psi}(\sigma) \ , \qquad (90a)$$

$$\epsilon = -\operatorname{sgn}[\partial\hat{\Phi}/\partial\sigma_n] \ . \qquad (90b)$$

In this new form, the main import of the Maximum Principle, (90b), is equivalent to a much simpler idea, namely that $\hat{\Phi}$ is a positive definite Liapunov function for the bang-bang control system (89) which (before chattering) is "stable" by virtue of having $-\hat{\Psi}$ as its negative Lie derivative. Solution of (90) is trivial and yields as the general solution of (84)–(86)

$$\Phi = \Phi_0(\sigma_1(x, \epsilon), \ldots, \sigma_{n-1}(x, \epsilon)) + \Phi_1(\sigma(x, \epsilon)) \ , \qquad (91a)$$

$$\Phi_1(\sigma) = -\epsilon\int_0^{\sigma_n} \Psi(h(\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \mu))\,d\mu \ ,$$

$$\epsilon^2 = 1 \ , \qquad (91b)$$

where $\Phi_0 = \Phi_0(\sigma_1, \ldots, \sigma_{n-1})$ is an arbitrary function of its n-1 arguments. Assuming $\Phi(x)$ continuous, any surface of discontinuity of $\epsilon(x)$ is constrained by the requirement that $\Phi_0(\sigma(x, -1)) + \Phi_1(\sigma(x, -1)) = \Phi_0(\sigma(x, +1)) + \Phi_1(\sigma(x, +1))$ for x on the surface.

Quadratic Performance Criteria

In [11] it was shown that if $\Psi = 1/2\ x\cdot Cx$ is a positive-definite quadratic form, then for x sufficiently near $x = 0$

$$\epsilon = \operatorname{sgn}[\sigma_0(x)] \ , \quad \sigma_0 = -(\sigma_n - \Phi_n(\sigma_1, \sigma_2, \ldots, \sigma_{n-1})) \qquad (92)$$

for a suitable function Φ_n. Also there exists a non-negative definite matrix B such that, in (91),

$$\Phi_0 = \frac{1}{2} h(\sigma_1, \ldots, \sigma_{n-1}, \Phi_n)\cdot Bh(\sigma_1, \ldots, \sigma_{n-1}, \Phi_n)$$
$$- \Phi_1(\sigma_1, \ldots, \sigma_{n-1}, \Phi_n) \ . \qquad (93)$$

Time-Optimal Control

When $\Psi \equiv 1$, the celebrated time-optimal problem is obtained. It will be shown here that, near $x = 0$,

$$\epsilon = \operatorname{sgn}[\sigma_0(x)] \ , \quad \sigma_0 = -\left(\sigma_1 - \Phi_2(\sigma_2, \ldots, \sigma_{n-1})\right) \ , \qquad (94a)$$

$$\sigma_i = \sigma_i(x, \hat{\epsilon}), \ \hat{\epsilon} = \hat{\epsilon}(x) \ , \ \hat{\epsilon}^2 = 1, \ (i = 1, 2, \ldots, n-1) \ , \qquad (94b)$$

$$(\Phi_2/||x||) \to 0 \quad \text{as} \quad ||x|| \to 0 \quad , \tag{94c}$$

where $\Phi_2(\sigma_2, \ldots, \sigma_{n-1})$ is an analytic function of fractional powers of its arguments, and where $\hat{\epsilon}(x)$ is characterized on and off the surface $\sigma_o(x) = 0$ by

$$\epsilon = -\hat{\epsilon} \quad , \quad \sigma_o \neq 0 \quad , \tag{94d}$$

$$\epsilon = \hat{\epsilon} \quad , \quad \sigma_o = 0 \quad . \tag{94e}$$

It is known [19], [18] that for linear controllable n^{th} order systems with real eigenvalues, time-optimal control can be effected by at most n-1 switches of a bang-bang control ϵ. For systems with complex eigenvalues, this result remains valid for initial conditions sufficiently near the origin in state space. Thus the state trajectory of (89) originates at x° with a specific ϵ, say $\epsilon_o = \epsilon(x^\circ)$. When x(t) crosses the switching surface, the system is governed by

$$\dot{x} = Ax - a\,\epsilon \tag{95}$$

during the next arc of the trajectory. Thus, the optimal switching surface for (89) is an integral of (95), which therefore must be some function of $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$.

The solution of (95) is

$$x(t) = e^{At}\,x^1 - \int_0^t \epsilon\, e^{A(t-\mu)}\, a d\mu \quad . \tag{96}$$

If x^1 is on the switching surface, then

$$0 = e^{At_{n-1}}\, x^1 \sum_{j=0}^{n-2} (-1)^{j+1}\, \epsilon \left[\int_{t_j}^{t_{j+1}} e^{A(t_{n-1}-\mu)} a d\mu\right], \tag{97}$$

where t_o is the time at which x(t) enters the surface, where $t_1, t_2, \ldots, t_{n-2}$ are the subsequent switching times, and t_{n-1} is the stopping time. Solving for x^1 and applying the convenient substitution,

$$\tau_j \stackrel{\Delta}{=} -t_{n-j-1} \quad , \qquad (j = 0, 1, \ldots, n-1) \quad , \tag{98}$$

yields

$$x^1 = \epsilon \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{(\nu+1)!} \sum_{j=0}^{n-2} (-1)^j \left[(-\tau_{n-j-2})^{\nu+1} - (-\tau_{n-j-1})^{\nu+1}\right] A^\nu a \quad . \tag{99}$$

Thus, taking $t_o \equiv \tau_{n-1} = 0$, the parametric form of the switching surface is

$$x = x(\tau) \stackrel{\Delta}{=} (-1)^{n-1}\, \epsilon \sum_{\nu=1}^{\infty} \hat{\varphi}_\nu(\tau)\, A^{\nu-1}\, a \quad , \tag{100}$$

$$\hat{\varphi}_\nu(\tau) \stackrel{\Delta}{=} \frac{1}{\nu!}\left[\tau_o^\nu + 2\sum_{j=1}^{n-2} (-1)^j\, \tau_j^\nu\right] \quad . \tag{101}$$

The tangent hyperplane at x = 0 is given by $q \cdot x = 0$, where q is the unit vector whose scalar product with (100) identically removes the terms in τ_o^ν, $\tau_1^\nu, \ldots, \tau_{n-2}^\nu$, ($\nu = 1, 2, \ldots, n-1$). By (7) it is clear that $q = b/||b||$ as claimed in (8). Thus it is seen without further calculation that the integral surface $\sigma_o = 0$ must be expressible in the form $\sigma_1 - \Phi_2(\sigma_2, \ldots, \sigma_{n-1}) = 0$, where $\Phi_2 = o(||x||)$.

The general properties (94e) of the auxiliary switching function $\hat{\epsilon} = \hat{\epsilon}(x)$ are obvious. The detailed procedure for calculating $\hat{\epsilon}(x)$ and $\Phi_2(\sigma_2, \ldots, \sigma_{n-1})$ is based upon simplification of (100) by means of (56a). First (52), (7), and repeated use of (49a), yield

$$\xi(s) = (-1)^{n-1}\epsilon \sum_{\nu=1}^{\infty} \hat{\varphi}_\nu(\tau)\left[s^\nu - \Delta(s)\sum_{k=1}^{\nu} (b\cdot(A^*)^{\nu-1} a)\, s^{j-k}\right]. \tag{102}$$

Then, with (69) and the Calculus of Residues, (56a) becomes

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)}\,\frac{\epsilon}{s} \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \left[(-1)^{n-1}\,\epsilon \sum_{\nu=1}^{\infty} \hat{\varphi}_\nu(\tau)\, s^\nu\right]^j ds \quad , \tag{103}$$

or equivalently, if $|s\hat{\xi}(s)| < 1$,

$$\sigma_j = \frac{1}{2\pi\sqrt{-1}} \oint_{|s|=\rho} \frac{s^{j-1}}{\Delta(s)}\,\frac{\epsilon}{s} \log(1 + \epsilon s\,\hat{\xi}(s))\, ds \quad , \tag{104a}$$

$$\hat{\xi}(s) \stackrel{\Delta}{=} (-1)^{n-1}\,\epsilon \sum_{\nu=1}^{\infty} \hat{\varphi}_\nu(\tau) s^\nu \quad . \tag{104b}$$

Defining $\hat{\omega}_j$, ($j = 1, 2, 3, \ldots$) by

$$(-1)^{n-1} \log(1 + \epsilon s\hat{\xi}(s)) \stackrel{\Delta}{=} \sum_{j=1}^{\infty} \hat{\omega}_j s^{j-1} \quad , \tag{105}$$

and proceeding as in Part IIC, the new parametric form

$$(-1)^{n-1}\,\epsilon\,\sigma_i = \sum_{\ell=0}^{\infty} \beta_\ell\, \hat{\omega}_{\ell+n-i+1}(\tau) \quad , \tag{106a}$$

$$\hat{\omega}_1 = \hat{\varphi}_1(\tau) ,$$

$$\hat{\omega}_\nu = \hat{\varphi}_\nu(\tau) + \frac{(-1)^{n-1}}{\nu} \sum_{m=1}^{\nu-1} m\,\hat{\omega}_m \hat{\varphi}_{\nu-m}(\tau) \quad ,$$

$$(\nu = 2, 3, 4, \ldots) \quad , \tag{106b}$$

can be derived. The function $\hat{\epsilon}$ is defined by solving (106a) for $i = 2, 3, \ldots, n$ [20] to get τ_i, ($i = 0, 1, 2, \ldots, n-2$) as functions of $\sigma_2, \ldots, \sigma_n$ and noting that τ_i are real and such that $\tau_o < \tau_1 < \ldots < \tau_{n-2} < 0$. Since the σ_i are integrals for ($i = 1, 2, \ldots, n-1$), they are unchanged by letting $\tau_{n-2} \to 0$. Now eliminate [20] the n-2 parameters $\tau_o < \tau_1 <, \ldots, < \tau_{n-3}$ between the (n-1) equations (106a), for $i = 1, 2, \ldots, n-1$, obtaining $\sigma_1 = \Phi_2(\sigma_2, \ldots, \sigma_{n-1})$ where Φ_2 is an analytic function of fractional powers of its arguments. Clearly $\sigma_o = \pm(\sigma_1 - \Phi_2)$ where the choice of + or - is unchanged by continuous variation of A or a. When $\Delta(s) = s^n$, as in (73)-(74) elementary arguments show that $\sigma_o = -\theta_1 + \ldots$ whence $\sigma_o = -(\sigma_1 - \Phi_2)$ in general.

References

1. Lur'e, A.I., Some Non-Linear Problems in the Theory of Automatic Control, London, Her Majesty's Stationery Office, 1957.

2. Letov, A.M., Stability in Nonlinear Control Systems, Princeton, Princeton University Press, 1961.

3. La Salle, J. and Lefschetz, S., Stability by Liapunov's Direct Method with Applications, New York, Academic Press, 1961.

4. Bass, R.W. and Mendelson, P., Aspects of General Control Theory, Air Force Office of Scientific Research Report No. AFOSR2754, August 1962.

5. Lewis, D.C. and Mendelson, P., "Contributions to the Theory of Optimal Control," Transactions of the American Mathematical Society, Vol. 110, No. 2, pp. 232-244, February, 1964.

6. Wonham, W.M. and Johnson, C.D., "Optimal Bang-Bang Control with Quadratic Performance Index," Transactions of the ASME-Journal of Basic Engineering, pp. 107-115, March, 1964.

7. Wonham, W.M. and Johnson, C.D., "On a Problem of Letov in Optimal Control," Preprints of Papers – 1964 JACC, pp. 317-325, June, 1964.

8. Wonham, W.M. and Johnson, C.D., "A Note on the Transformation to Canonical (Phase-Variable) Form," IEEE Transactions on Automatic Control, pp. 312-313, July, 1964.

9. Kalman, R.E., "When is a Linear Control System Optimal?," Transactions of the ASME Journal of Basic Engineering, pp. 51-60, March, 1964.

10. Bass, R.W. and Gura, I., "High Order System Design via State Space Considerations," Proceedings of 1965 JACC, pp. 311-318.

11. Bass, R.W. and Webber, R.F., "On Synthesis of Optimal Bang-Bang Feedback Control Systems with Quadratic Performance Index," Proceedings of 1965 JACC, pp. 213-219.

12. Mufti, I.H., "On the Reduction of a System to Canonial (Phase Variable) Form," IEEE Transactions on Automatic Control, pp. 206-207, April, 1965.

13. Bass, R.W., Gura, I., and Webber, R.F., et al, Final Report, Minimax Attitude Control of Aeroballistic Vehicles, NASA Contract NAS8-11421, Marshall Space Flight Center, Huntsville, Alabama, July, 1965.

14. Bass, R.W. and Webber, R.F., "Optimal Nonlinear Feedback Control Derived from Quartic and Higher-Order Performance Criteria," (submitted to 1966 IFAC).

15. Zadeh, L.A. and Desoer, C.A., Linear System Theory, New York, McGraw-Hill Book Company, Inc., pp. 303 ff., 1963.

16. Hartman, P., Ordinary Differential Equations, John Wiley, New York, 1964.

17. Andre, J. and Seibert, P., "Piecewise Continuous Differential Equations," Boletin de la Sociedad Mathematica Mexicana, 1961, pp. 242-245.

18. Pontriagin, L.S., Boltianski, V.G., Gamkrelidze, R.V., and Mischenko, E.F., The Mathematical Theory of Optimal Processes Interscience, New York, 1962.

19. Bellman, R., Glicksberg I., and Gross, O.A., Some Aspects of Mathematical Theory of Control Processes, The Rand Corp., Project Rand R-313, January, 1958.

20. Osgood, W.F., Lehrbuch der Funktionentheorie, 2nd ed. (1929, Vol II_1, p. 132.

Acknowledgment – This research was performed under NASA Contract NAS8-11421.

N66-12129

ERRATA

NASA CR-350

MINIMAX ATTITUDE CONTROL OF AEROBALLISTIC LAUNCH VEHICLES

December 1965

Page No.	Line	
5-8	2	Delete MATPWR,
5-8	11	Add ", ELINV1(5, 5)" after ", TEMP(5)."
5-10	24	Change "ELINV" to read "ELINV1."
5-22	22	Change to read "IF(NEVN. EQ. O)COEFF(N+1)=1."
5-22	36	Add "+PART 2" to end of line.

NASA-Langley, 1966

Issued 10-26-66